Reap What You Sow: Agricultural Technology, Urbanization and Structural Change

Danny McGowan; Vasilakis, Chrysovalantis

Research Policy

DOI:
10.1016/j.respol.2019.05.003

Published: 01/11/2019

Peer reviewed version

Reap What You Sow: Agricultural Technology, Urbanization and Structural Change

Danny McGowan* and Chrysovalantis Vasilakis†‡

May 8, 2019

Abstract

This paper studies how productivity-enhancing agricultural technology affects urbanization by provoking structural change. We investigate these issues using a natural experiment in the United States. The results show that technologies which improve crop productivity lead to a less urbanized economy as economic activity relocates from manufacturing and services towards agriculture. The effects are highly persistent and are driven by the technology increasing agricultural labor demand. Our findings highlight the potentially unintended, disruptive force of innovative technologies.

Keywords: agricultural technology, productivity, urbanization, structural change

JEL Codes: D22, D24, L16, Q11

*Corresponding author. University of Birmingham. Email: d.mcgowan@bham.ac.uk. Tel: +44 (0)121 414 6678.

†Bangor University and IZA. Email: c.vasilakis@bangor.ac.uk.

‡We thank the editor, Maria Savona, two anonymous referees, Mirko Draca, Frederic Docquier, Oded Galor, Paola Guiliano, Richard Kneller, William Pariente, Luigi Pascali, and seminar participants at the University of Louvain and the University of Warwick for helpful comments and suggestions. We are grateful to Leah Boustan for sharing data. Samuel Bottom provided excellent research assistance.

1 Introduction

Do innovations provoke societal changes? If so, to what extent? Innovations continuously disrupt existing industries, leading to changes in sectoral employment shares and where people live (Schumpeter, 1942; Fagerberg and Verspagen, 2002; Weber and Rohracher, 2012). Key to this process are new productivity-enhancing technologies (Quatraro, 2009). For example, new manufacturing technologies triggered a reallocation of employment from agriculture to manufacturing and the growth of cities during the 19$^{\text{th}}$ century. Recently, the rise of the service sector and growth of urban areas has been underpinned by new information and communication technologies (Bartel et al., 2007; Holm and Ostergaard, 2015).

Theoretical models formalize these ideas and suggest agricultural innovations are particularly important. Early contributions suggest that agricultural technologies lead to industrialization and the growth of cities by releasing labor from rural areas, triggering migration to urban areas as workers search for employment in manufacturing and service industries (Clark, 1940; Rosenstein-Rodan, 1943; Kuznets, 1957; Gollin et al., 2002). However, evidence also suggests that improvements in agricultural productivity may crowd out non-agricultural sectors (Mokyr, 1976; Field, 1978; Wright, 1979). Matsuyama (1992) illustrates this idea in a model where improving crop productivity leads non-agricultural sectors to contract in an open economy.

Related research outlines theoretical mechanisms through which agricultural productivity shocks influence urbanization by provoking structural change. In these models, a region comprises an agricultural, manufacturing and service sector which rely on capital, labor and land inputs to produce output (Hornbeck and Keskin, 2015). Labor is supplied by workers within the region who may live either in rural or urban areas. Workers are mobile across sectors but commuting costs exclude the possibility that an individual can live in the rural area and work in the urban area, and vice versa (Roback, 1982).

Technologies that raise agricultural productivity increase agricultural demand for land and labor as farms require more workers to process the additional output leading to higher wages (Foster and Rosenzweig, 2004; Hornbeck and Keskin, 2015). In addition, as workers' utility depends on wages and housing costs, higher demand for land raises workers' housing costs, leading them to demand higher wages in compensation (Hornbeck and Keskin, 2015). As the cost of land and labor inputs increase, the manufacturing and service sectors contract. However, for non-tradable services, which are consumed locally, these effects may be offset by increased demand from the agricultural sector (Hornbeck and Keskin, 2015). These productivity-induced changes in the structure of the economy also influence urbanization. As agriculture is concentrated in the rural area, improvements to agricultural productivity lead to lower urbanization rates as workers move from manufacturing and services to agriculture in search of employment (Hornbeck and Keskin, 2015).

In this paper, we explore these predictions by studying the effects of a breakthrough agricultural innovation: hybrid corn. Hybrid corn seed dramatically increased crop productivity (output per acre) in US counties suited to growing corn. Invention of the hybrid technology was exogenous with respect to the economic outcomes we study which allows us to use difference-in-difference estimations to compare the cross-time evolution of urbanization and economic structure within the treatment (counties suited to growing corn) and control (counties unsuited to growing corn) groups.

The empirical results provide strong support for the view that agricultural productivity is a key determinant of urbanization and economic structure. We find that the hybrid technology caused a reallocation of economic activity from manufacturing and services towards agriculture, leading to a less urbanized economy. For example, agriculture's share of employment increases by approximately 5% whereas the manufacturing and tradable service sector employment shares fall by 32% and 6%, respectively. In contrast, non-tradable services are unaffected by the productivity shock. Consistent with these patterns, we also find that the treatment group experiences a reduction in urbanization. In terms of magnitude, we find that the rate of urbanization fell by approximately 33% within the average treated county.

We conduct several robustness checks to ensure our findings are not driven by confounding events. We show that differences in the severity of the Great Depression across regions, developments within the manufacturing sector and differences in the intensity of the Dust Bowl do not confound our results. Further tests rule out that the results stem from subsidies provided through the New Deal program, other agricultural productivity shocks and the rapid development of the western US during the sample period. Falsification exercises verify that the results are indeed driven by the adoption of hybrid corn seeds.

Our paper relates to three strands of literature. One area of research studies the role of agricultural productivity in provoking a transition from a rural to industrialized economy. Bustos et al. (2016) show that the effects of agricultural productivity on structural change depends on whether crop productivity or agricultural labor productivity increases. Improvements to crop productivity lead to a more agrarian economy at the expense of the tradable goods sector while increasing agricultural labor productivity has the opposite effect. However, they find the relationship between crop productivity and the size of the service sector to be ambiguous. Recent empirical studies also report mixed evidence. For example, Jedwab (2013) finds that agricultural development provokes urbanization and increases the size of the service sector whereas Hornbeck and Keskin (2015) find no effect on either outcome.

In contrast, we find that improving crop productivity causes a relocation of the workforce not only from manufacturing but from tradable services as well. A potential explanation for the differences between our findings and Hornbeck and Keskin (2015) is that during our sample period labor mobility is more restricted. This would produce larger urbanization and structural change reactions to agricultural productivity shocks. For example, where migration costs increase with geographic distance, workers may be disproportionately drawn from within the county. Productivity shocks that raise agricultural labor demand therefore create more intense competition for local labor which is not offset by immigration from outside counties. Non-agricultural sectors therefore experience a more pronounced rise in labor costs compared to a setting with fewer migration impediments.

Another potential explanation is that agriculture is less mechanized during our sample period compared to the setting in Hornbeck and Keskin (2015). In a less mechanized environment farms cannot respond to productivity shocks by employing more capital. Hence, innovations that increase output lead to higher agricultural labor demand that spillover to non-agricultural sectors.[1]

Jedwab (2013) documents that urbanization without structural transformation is possible if wealth is created in rural areas and spent on non-tradable services in urban areas. The contrast between his findings and ours may be due to the non-tradable service sector being distributed across rural and urban areas within our sample. In this setting, improvements to agricultural productivity do not disproportionately create higher demand for urban non-tradable services. Jedwab (2013) also highlights that deurbanization may occur where regions have a comparative advantage in agriculture. This may be the case in our setting.[2]

Finally, our research contributes to the innovation literature. Schumpeter (1942), Kuznets (1972) and Audretsch and Klepper (2000) recognize that innovations alter the structure of the economy by creating jobs in new and existing industries while rendering others obsolete. At the heart of this process are technologies that improve firm productivity and spur firm growth (Brynjolfsson and Hitt, 2003; Aw et al., 2008; Boothby et al., 2010; Syverson, 2011; Criscuolo et al., 2012; Coad et al., 2016).[3] Since the Industrial Revolution, cycles of innovation-driven sectoral reallocation have coincided with shifting patterns of urbanization. Dolata (2009) highlights that whereas the determinants of innovation have received great attention, understanding how socioeconomic structures and economic actors

[1] A separate possibility is that differences in the productivity effects of the technologies plays a role. For example, hybrid corn has a large unambiguous effect on crop productivity. In contrast, Hornbeck and Keskin (2014) find the Ogallala aquifer technology has no effect on crop productivity outside drought periods. Rather it led farmers to specialize in producing water-intensive commodities. The Ogallala aquifer may therefore have a more muted effect on urbanization because it has a smaller impact on crop productivity and agricultural labor demand relative to hybrid corn.

[2] A related strand of literature focuses more broadly on the role of agricultural productivity in determining structural change. Key references include Gollin et al. (2002) who find that improvements in productivity within agriculture lead to a more industrialized economy and Foster and Rosenzweig (2004) who obtain the opposite result.

[3] A related strand of literature documents the productivity-enhancing effects of patents and R&D investments. See Hall (2011) and Hall et al. (2011) for reviews.

change under the influence of new technologies is an understudied area. Our paper helps fill this gap.

The paper proceeds in the following order. Section 2 provides details about the data. Background information on our economic laboratory is provided in Section 3. We outline our identification strategy, present econometric results and robustness tests in Section 4. Finally, we draw conclusions and discuss policy implications in Section 5.

2 Data and Summary Statistics

We use two separate data sets to study the effect of hybrid corn seeds on, 1) crop productivity, and 2) urbanization and structural change.

2.1 Agricultural Productivity Data

We retrieve information on crop productivity from the National Agricultural Statistics Service (NASS), the statistics branch of the US Department of Agriculture. This source provides annual county-level crop productivity data for a variety of field crops over the period 1930 to 1975. We therefore have information on crop productivity in county c for crop i during year t. Crop productivity corresponds to output per acre and is measured in bushels per acre.[4] The NASS also provides information on the number of acres planted and total output (measured in bushels) of each crop in each county-year. To sharpen identification we use a balanced panel comprising only counties that produce both corn and wheat. That is, we compare the evolution of crop productivity between the corn and wheat industries within the same county. This feature obviates confounding factors as both crops are subject to the same climatic conditions and local economic shocks that may affect crop productivity more generally. Our main tests focus on the period 1930 to 1940. In total this provides a sample containing 4,070 observations.

[4]A bushel is approximately equal to 25 kilograms.

2.2 County-Level Data

Our main data source for the urbanization and structural change tests is the US Census. This database provides county-level information on urbanization rates, agricultural and manufacturing variables for the years 1920, 1930 and 1940.[5] We compute urbanization rates as the ratio of people living in urban areas to total population in the county. We measure structural change using sectoral employment shares. For example, the agricultural employment share (AG share) is the ratio of agricultural employment to total employment within county c during year t. The manufacturing share (M share) is the ratio of manufacturing employment to total employment within county c during year t. Information on the number of workers employed in the tradable and non-tradable sectors within each county-year is taken from the IPUMS database. We compute the tradabe ($\mathrm{S^T}$ share) and non-tradable ($\mathrm{S^N}$ share) service share as the ratio of tradable and non-tradable service sector employment to total employment, respectively.

In addition, the Census data provide county-level information on population density (population per square kilometer), manufacturing productivity (the ratio of manufacturing goods value to manufacturing employees), manufacturing establishments per 1,000 population, and average farm size (acres per farm). We also calculate the price of land per square mile (the ratio of farmland value to square miles) and manufacturing wages (the ratio of annual manufacturing wages to manufacturing employment) in each county-year.

To capture business-cycle fluctuations we construct the state-level unemployment rate data using the IPUMS database and the method outlined by Boustan et al. (2010).[6] We also retrieve county-level information on the fertility rate (net births and deaths per 1,000 population) in each county from the US Census. To capture the extent to which a county was affected by the Dust Bowl, we retrieve county-level erosion data from Hornbeck (2012).

[5]We restrict the sample to these years because the variables available through the US Census change over time. Using these years ensures that we have comparable measures for the key outcomes we are interested in across time.

[6]The results are unchanged when we use county-level unemployment rate data from the Census. However, we prefer the state-level variable on the grounds it constructs unemployment rates using the labor force which is not available in in county-level Census data.

Information on payments to farmers through the Agricultural Adjustment Act of 1933 is taken from Fishback (2005).

Finally, to measure the size of the agricultural sector in each county, we calculate agricultural output per capita. We first retrieve information on the physical quantity of each agricultural commodity (corn, cotton, hogs, wheat etc.) produced in each county-year from the NASS. The NASS also provides data on the price per physical unit of each commodity. Using the price and quantity data we calculate output in each commodity (price times quantity) and sum across all commodities to give agricultural output in the county. We then weight this by population to obtain agricultural output per capita.

2.3 Corn Suitability

The empirical strategy exploits the fact that some regions of the US were affected by the invention of hybrid corn seeds whereas others were not owing to their unsuitability to produce corn. The Food and Agriculture Organization's (FAO) Global Agro-Ecological Zones (GAEZ) database uses agronomic models and high resolution data on geographic characteristics such as soil quality, topography, elevation and climatic conditions to calculate the suitability of an area to growing corn. This source provides a suitability index ranging between 0 (unsuitable) and 100 (highly suitable) for each cell on the Earth's surface.[7] We aggregate the grid cells to the county-level to obtain a measure of corn suitability for each county.

Importantly, the suitability index is available for all cells regardless of whether corn is actually grown in a given location. Moreover, the index is not based on specific contemporary or historical technologies (e.g. traditional or hybrid corn seeds). The key determinants of suitability are local climatic and topographical characteristics as well as soil conditions that have not changed much over time. See section 6 in Fischer et al. (2012) for further details.

[7] A cell is defined as a 5 arc-minute, approximately equivalent to a 10km^2 grid. Other economic studies which use the FAO-GAEZ database to construct suitability measures include Costinot et al. (2015) and Nunn and Qian (2011).

A concern is the extent to which the FAO-GAEZ suitability measure is an accurate indicator of suitability during the 1930s and 1940s given it is constructed by researchers to inform contemporary individuals and governments. The construction of the suitability measure does not give any obvious cause for concern. In fact, the suitability index should be a good proxy for historical conditions because they are primarily based on climatic characteristics such as temperature, humidity, length of days, sunlight, topography and rainfall that have not changed significantly since the period of our study.[8]

2.4 Summary Statistics

Table 1 provides a summary of the variables we use in the empirical analyses. Panel A tabulates the variables we use to examine the effect of hybrid corn seed on crop productivity whereas Panel B provides a summary of the variables used in the urbanization and structural change tests. The urbanization rate is approximately 21% in the average county while agriculture and manufacturing account for around 46% and 5% of employment, respectively. 9% of counties belong to the treatment group (outlined in Section 4.1).

[Insert Table 1]

3 Institutional Details

Corn is one of the largest industries within the US agriculture sector, and typically accounts for between 30% and 40% of agricultural output. Corn plants are made up of a leafy stalk that produces ears containing kernels (seeds). The kernels may be used to produce food, livestock feed or future generations of corn seed. The more kernels a plant produces, the higher its yield and the greater is crop productivity.

[Insert Figure 1]

[8] Moreover, in constructing our suitability measure, we intentionally used the FAO-GAEZ suitability measure that assumes rain-fed conditions and medium input usage to approximate historical conditions as closely as possible and avoid measurement error from changes over time in irrigation intensity and technologies. Nunn and Qian (2011) also follow this approach.

Until the early 1930s corn farmers only had access to traditional (single-cross) seeds. Seed companies produced traditional seeds by randomly mating two corn plants and harvesting their kernels for the next growing season.[9] An acre planted with traditional seeds yielded, on average, 26 bushels per acre. Figure 1 provides an illustration of historical crop productivity in the corn industry. It is apparent from the figure that before 1935 corn yields per acre were essentially constant with some fluctuations around the mean due to weather shocks. The absence of crop productivity improvements during this period was due to the fact that while seed producers could to some extent manipulate corn traits by selecting the best plants, selection was ineffective in raising yields (Crow, 1998; Troyer, 2009).[10]

Beginning in the early 1900s geneticists discovered that crossing inbred corn plants could occasionally produce higher yields compared to traditional varieties (Shull, 1908, 1909, 1910, 1911). However, inbred plants' resistance to disease was often poor and most never grew to maturity. Development of commercially viable inbred varieties followed a trial and error approach at universities and agricultural research stations. Hybrid corn seeds were finally discovered in 1918 and became commercially available in 1934 following several years of refinement (Griliches, 1957; USDA, 1962; Skinner and Staiger, 2007; Troyer, 2009).[11] Hybrids produced vastly more kernels relative to traditional varieties as they combined the yield advantages of inbred plants with the disease resistance of traditional plants.[12]

[9] Plants were randomly mated through wind-borne pollination in fields owned by a seed company. Following successful pollination a plant is fertilized and the reproduction process begins leading to the development of kernels. Seed producers then harvest the plants and store the kernels to be sold as seed the next year.

[10] Corn reproduces sexually each year. This process randomly selects half the genes from a given plant to propagate the next generation. Consequently, mass production of seeds with desirable genes (such as yield) can be lost in subsequent generations as high- and low-yielding plants are randomly mated. Selection of high yielding traditional plants was therefore not viable. See the Online Appendix for details.

[11] Hybrid corn is a plant produced by mating two inbred corn plants. Experimental trials in the early 1930s found double-cross hybrids - crossing two inbred plants and crossing that hybrid with the hybrid of two other inbred plants - to reliably produce high yields.

[12] Planting hybrid and traditional varieties in close proximity has no effect on the productivity of either type of plant. This is because the plant's genetic make-up (that is, the number of kernels it produces) is determined solely by the traits in its seed. Instead, if a hybrid and traditional plant were to mate, it would only determine the genetic composition of the next generation of seeds. The current generation's genetic make-up would be unaffected as this depends on the seed from which the plant germinated. Given that farmers purchase seed from seed companies for one growing season at a time, proximity of

[Insert Figure 2]

Figure 2 shows that hybrid corn seed was rapidly adopted following its invention. Whereas the new technology was unused during the early 1930s, after becoming available the incidence of hybrid corn seed quickly increases after 1935 and accounts for in excess of 30% of planted acres by 1940. These patterns are consistent with a number of historical accounts. For example, Griliches (1957) reports that hybrid corn became available in the mid 1930s and was rapidly adopted by farmers. In addition, he shows that the rate of adoption was considerably faster in Midwestern states and reached 80% of planted acres by 1940. Dixon (1980) and Skinner and Staiger (2007) report similar evidence. The rapid adoption of the technology was due to academic evidence from the Iowa Corn Yield Test that demonstrated the productivity advantages of hybrid corn (Troyer, 2009).

3.1 Nature of the Research Process

Early attempts to develop more productive corn seeds were inspired by theoretical and empirical insights provided by Darwin (1859, 1876). This framework identified the negative effects of inbreeding and led researchers to investigate whether crossing seed varieties improved yields. Guided by these insights, from at least 1846 farmers and researchers began searching for more productive corn seeds by crossing different varieties of corn seed (Troyer, 2009). This process involved four steps.

1. Select seeds from corn plants with desirable properties. For example, in developing the Reid Yellow Dent variety Robert and James Reid selected seeds from corn plants that produced higher-yielding offspring and plants with ideal agronomic properties such as medium-sized ears, bright yellow colors, and solid, deep and relatively smooth grains.

2. Researchers then cross-pollinated plants belonging to different varieties. In the example above, the Reids cross-pollinated plants belonging to the Gordon Hopkins and

the two varieties had no effect on the productivity of each variety during the current growing season.

Little Yellow varieties. Cross-pollination was done by hand. A researcher collected pollen from one plant in a bag, then place the flowers from another plant variety in the bag, and shook it to ensure the pollen settled on the plant's stigma.

3. After cross-pollinating a plant, researchers harvested the seeds produced and plant these single-cross seeds. They then evaluated whether the new variety produced higher yields than their parents and other available varieties.

4. It typically took 15 years of further selection and refinement before a seed variety was suitable for commercial sale (Wallace, 1955; Troyer, 2009).

In contrast to single-cross varieties, research into developing hybrid corn was almost exclusively conducted by academics working at agricultural research stations and universities.[13] Troyer (2009) reports hybrid corn researchers were motivated by theoretical insights provided by Shull (1908, 1909, 1910, 1911). While experiments followed the same approach as outlined in steps 1 to 4 above, a notable difference was that researchers created inbred corn plants by pollinating a plant using its own pollen. They then cross-pollinated two inbred varieties.[14]

Throughout this process collaboration and interactions between researchers were pivotal. Troyer (2009) reports exchanges between William Beal and Charles Darwin, Edward East, George Shull, and Donald Jones, and between researchers working at the University of Illinois and the Connecticut Experiment Station. This research network was essential in developing the hybrid concept and the search for inbred and hybrid corn seeds. The Online Appendix contains details of the experiments undertaken.

Private firms chose not to invest in developing hybrid corn seeds because the research and development costs were so high that expected returns were negative (Wallace and

[13]Especially important were the University of Illinois, Harvard University, Cornell University, and the Connecticut Experiment Station (Troyer, 2009).

[14]Initially, researchers focused on crossing non-inbred varieties based on evidence reported by Darwin (1859, 1876) that showed inbred plants had a number of unattractive properties (relative to their parents they are smaller, produce lower yields, are less disease resistant, etc.). The idea of crossing inbred plants (a necessary requirement for hybrid corn seed) was not known. Research into inbred corn plants began in the early 1900s (for example, Edward East in 1905 and H.K. Hayes in 1909).

Bressman, 1923, 1949). Public funding for universities and agricultural research stations was therefore essential in establishing the research clusters that led the search for hybrid corn. These funds were mainly provided through the Morrill Act of 1862 and by the Harding Administration in the early 1920s.

3.2 The Effect of Hybrid Corn Seed on Crop Productivity

Figure 1 illustrates the rapid increase in corn productivity after the introduction of hybrid corn seed. To pin down a precise estimate of the crop productivity effects of the hybrid technology we turn to regression analysis. We use difference-in-difference estimations which compare the evolution of crop productivity within the corn industry relative to productivity in the wheat industry.[15] Unlike corn, there were no coinciding technological developments within the wheat industry. Using the NASS productivity data we estimate the equation

$$yield_{ict} = \alpha_c + \beta_1 Corn_{ic} + \beta_2 Post_t + \beta_3 Corn_{ic} * Post_t + \delta X_{ict} + \gamma_t + \varepsilon_{ict}, \qquad (1)$$

where $yield_{ict}$ is output per acre in industry i of county c at time t (in natural logarithms); $Corn_{ic}$ is a dummy variable equal to 1 if the observation is from the corn industry, 0 otherwise; $Post_t$ is a dummy equal to 1 for the years 1935 onward, 0 otherwise; X_{ict} is a vector of control variables; α_c and γ_t are county and year fixed effects, respectively; ε_{ict} is the error term. In line with Bertrand et al. (2004) we cluster the standard errors at the county level.[16]

[Insert Table 2] [Insert Figure 3]

[15] Wheat is chosen as the control group because it is a close genetic relative of corn, is grown in similar areas, is planted and harvested at the same times, and relies on similar production methods. Wheat productivity therefore closely mirrors corn productivity. This allows us to construct a counterfactual of what would have happened to corn productivity in the absence of the new technology.

[16] This specification uses a balanced panel to ensure the results are not driven by counties entering or exiting the data set. That is, it includes only observations from counties that grow both crops in all years between 1930 and 1940. The results are unchanged when we use an unbalanced panel that includes observations of counties that begin producing either crop or drop out of the sample by stopping production.

Estimates of equation (1) are provided in Table 2. Column 1 of Table 2 reports the results of a simple specification that excludes control variables.[17] The coefficient on the corn variable indicates that within the average county crop productivity is approximately 63% higher in the corn industry relative to the wheat industry. The Corn-Post interaction coefficient is both positive and highly statistically significant. This shows that hybrid seeds caused a 19% increase in crop productivity relative to the counterfactual.[18]

To obtain more detailed inferences into the productivity dynamics, and inspect whether there were any anticipation effects, we examine the effects year by year. We estimate the equation

$$\begin{aligned} yield_{ict} = \alpha_c + \beta_1 Corn_{ic} + \phi_1 D1931_t + \phi_2 D1932_t + \cdots + \phi_1 D1940_t \\ + \varphi_1 Corn_{ic} * D1931_t + \varphi_2 Corn_{ic} * D1932_t + \cdots + \varphi_{10} Corn_{ic} * D1940_t + \delta X_{ict} + \varepsilon_{ict}, \end{aligned} \tag{2}$$

where all variables are defined as in equation (1) except $D1931_t$, $D1932_t$, ... , $D1940_t$ which denote year dummies for 1931, 1932, ... , and 1940, respectively. This approach provides insights into how much higher or lower corn productivity is relative to wheat productivity during year t.

We report estimates of equation (2) in column 2 of Table 2. To aid interpretation we present the interaction coefficient estimates and 95% confidence intervals for each year in Figure 3. For each year before 1935 there are no significant differences in productivity growth since 1920 between the two crops. In every instance the confidence intervals bisect zero. However, productivity begins to diverge following the invention of hybrid corn seed. In 1935, for example, corn productivity is significantly higher relative to wheat productivity. This is also the case for the remaining years. These pieces of evidence show that there were no anticipation effects and that hybrid corn seed triggered substantial

[17] The corn dummy variable captures the industry fixed effect.

[18] The productivity effects are consistent with experimental evidence. For example, using randomized control trials and traditional and hybrid corn seeds from 1935, Russell (1974) documents that under laboratory conditions hybrid corn plants produce between 60% and 100% more output compared to traditional seed varieties. This is driven by the genetic traits of hybrids that increase crop productivity.

improvements in corn productivity.

Next, we test the robustness of this finding by including a vector of control variables to ensure the productivity gains are not driven by confounding factors. For example, there may exist differences in initial conditions between counties that differentially affect crop productivity subsequently. We therefore include a set of interactions between a vector of initial conditions and the Post dummy variable. In column 3 of Table 2 we find that counties with higher manufacturing productivity in 1930 tended to have significantly lower crop productivity post 1930. The interactions between the Post dummy variable and the 1930 values of population density, the unemployment rate and fertility are all statistically insignificant. However, our key finding is unaffected. We continue to find that hybrid corn seed caused an economically large and statistically significant increase in corn productivity relative to the counterfactual.

To ensure the productivity effects are not driven by unobservable county-specific factors that differentially affect productivity across industries, we augment equation (1) with a set of county-industry fixed effects. The results in column 4 of Table 2 remain very similar to before.

We next conduct a validation check to ensure the productivity gains are driven by the adoption of hybrid corn seed and not time shocks. We interact the corn dummy variable with the annual share of acres planted using hybrid corn seed reported by USDA (1962) and include the interaction in equation (1). Intuitively, one would expect the corn-hybrid interaction coefficient to be positive and statistically significant: as the hybrid technology becomes more prevalent corn productivity increases. Indeed, this is what we observe in the estimates reported in column 5 of Table 2. A 1 percentage point increase in the share of acres planted with hybrid corn seed causes a 1.1% increase in corn productivity. In view of the hybrid adoption rates reported previously, this implies that hybrid corn seed increased productivity within the corn sector by approximately 33% between 1930 and 1940.

Finally, we examine whether the effect of hybrid corn on crop productivity is present over the long run and take steps to ensure the estimates are not driven by low corn productivity during the early 1930s. In column 6 of Table 2 we report estimates of equation (2) using a sample from 1909 to 1975. We continue to find that hybrid corn seed caused a statistically significant increase in corn productivity. Economically, the magnitude of the average treatment effect is somewhat larger compared to before, equivalent to a 53% increase relative to the counterfactual. This likely reflects that productivity across the corn industry increased as hybrid corn was adopted by a greater share of farmers. These tests also help to rule out that the productivity effects of hybrid corn seed are not driven by low productivity within the corn industry during the early 1930s. If this was the case, including a longer pre-treatment period ensures we capture such effects. However, the evidence in Figure 1 indicates that this is unlikely to drive our inferences. Before 1935 corn productivity is fairly constant across years and there is no evidence that it falls or is artificially low during the early 1930s.

Together the evidence in Table 2 demonstrates that hybrid corn seed caused a significant, persistent increase in corn productivity.

4 Empirical Analysis

In this section, we first outline the identification strategy, then present estimation results and discuss the findings, and conduct robustness checks.

4.1 Identification Strategy

To isolate causality our empirical strategy exploits two sources of exogenous variation. First, time series variation in crop productivity comes through the invention of hybrid corn seed. Second, we leverage time-invariant cross-sectional differences in counties' suitability to cultivate corn.

Only some counties are suited to growing corn. We assign counties to the treatment (control) group based on whether the county suitability index is greater or equal to (below)

70: equivalent to good suitability.[19] This choice of threshold is based on econometric evidence reported in Appendix Table A1. We use probit models to examine how well the suitability index explains corn cultivation. We implement this test by estimating

$$p_{ct} = \alpha + \beta_1 T_c + \beta_2 T50_c + \beta_3 T30_c + \delta X_{ct} + \gamma_t + \varepsilon_{ct} \tag{3}$$

where p_{ct} is a dummy variable equal to 1 if corn is planted in county c during year t, 0 otherwise; T_c is equal to 1 if the suitability index is 70 or above, 0 otherwise; $T50_c$ is equal to 1 if the suitability index is greater than or equal to 50 but less than 70, 0 otherwise; $T30_c$ is equal to 1 if the suitability index is greater than or equal to 30 but less than 50, 0 otherwise; γ_t are year effects; ε_{ct} is the error term.

The results in Appendix Table A1 show that a suitability index value of 70 or above significantly increases the probability that corn is grown in a county by 0.73%. The marginal effect is much weaker for $T50_c$ although the coefficient is also positive and statistically significant at conventional levels. However, the $T30_c$ variable is insignificant indicating that suitability values below 50 do not correlate with corn cultivation. We therefore conservatively define the treatment group as counties with suitability values above 70 although in unreported regressions our findings are robust to using a threshold of 50. Figure 4 shows the location of counties that are in the treatment and control groups.

[Insert Figure 4]

Our main tests revolve around difference-in-difference estimations of the equation

$$y_{ct} = \alpha_c + \beta Post_t + \varphi T_c * Post_t + \delta X_{ct} + \varepsilon_{ct}, \tag{4}$$

where y_{ct} is the outcome of interest in county c at time t; $Post_t$ is a dummy variable equal to 1 if an observation is from 1940, 0 otherwise; T_c is a dummy variable equal to 1 if the

[19] Corn accounts for a large share of agricultural activity within the treatment group. For example, corn acreage accounts for approximately 35% of planted acres in the average treatment-group county. Hence, the increase in crop productivity improved productivity within the agricultural sector more generally.

county is in the treatment group, 0 otherwise; $Post_t$ is a dummy variable equal to 1 in the year 1940, 0 in 1930; X_{ct} is a vector of control variables; α_c represent county fixed effects; and ε_{ct} is the error term.[20] Because T_c is a time invariant county specific characteristic, it is captured by the county fixed effects. As there are only two years in the sample, the year effects are captured by $Post_t$. To adjust for spatial correlation in the error term we follow the approach developed by Conley (1999) and used by Ashraf and Galor (2013).

Central to the internal validity of our research design is the exogeneity of hybrid corn. Wallace (1955), Crow (1998) and Troyer (2009) are explicit in their assessment that the invention of hybrid corn seed was driven by curiosity among agricultural researchers at universities and research stations. The discovery process was haphazard and followed a trial-and-error approach owing to geneticists' limited understanding of hybrid plants. This is exemplified by the discovery of hybrid corn seed in 1918 by Donald Jones. While conducting research at the Connecticut Experiment Station, Jones crossed inbred varieties that he did not realize were inbred. The seeds had been in storage since 1905 following experiments into inbreeding by other researchers. Hybrids did not become available immediately because seeds required refinement and substantial testing before they would grow reliably.[21] Wallace (1955) reports that this process took approximately 15 years for all corn varieties, including hybrids. Almost all of this research took place at agricultural research stations and universities (Johnson, 1957). During the mid-1930s academic researchers discovered double-cross hybrids that improved on Jones' discoveries by producing reliably high yields. This explains why hybrids first emerged in the mid-1930s (Troyer, 2009).

The invention of hybrid corn seeds was not driven by urbanization rates or US counties' economic structures. Rather it was due to luck in an experiment that predates our sample period. The interaction term is therefore uncorrelated with the error term in equation (4). Moreover, endogeneity does not arise through reverse causality. That would imply that urbanization and economic structure motivated the invention of hybrid seeds which was

[20] Because there are only two years in the sample, the year fixed effects are captured by $Post_t$.

[21] For example, Jones' single-cross hybrid lacked diseas resistance traits and often failed to grow to maturity (Crow, 1998).

not the case. See the Online Appendix for details.[22]

4.2 Diagnostic Tests

Identifying the effect of a treatment is more convincing where the treatment and control groups are similar before treatment occurs. We therefore use t-tests to examine the similarity of the treatment and control groups in 1930 and report the results in Table 3. We find no significant differences in the urbanization rate, the employment share of agriculture, manufacturing or services, population density, unemployment, manufacturing productivity, fertility, the number of manufacturing establishments per capita, manufacturing wages, and land prices between the treatment and control group. This finding suggests that the groups are indeed comparable ex ante, and the control group represents a good counterfactual for the treatment group.

[Insert Table 3] [Insert Table 4]

4.3 Results and Discussion

Table 4 reports estimates of equation (4). We begin by examining the impact of the hybrid technology on urbanization rates. In column 1 of Table 4 the average treatment effect is estimated to be -0.3955, indicating that urbanization rate fell by approximately 33% within the treatment group, relative to the counterfactual. The coefficient estimate is statistically significant at the 5% level. Hence, the hybrid technology provoked a substantial decrease in the rate of urbanization.[23]

Next, we test the sensitivity of this result to including of a vector of control variables in equation (4). The choice of controls is motivated by recent contributions to the literature on the determinants of urbanization that emphasize the role of initial population

[22] To examine whether conditions within corn-growing counties triggered the invention of hybrid corn we follow the procedure outlined by Danisewicz et al. (2017) to ascertain whether this was the case. The results of this test are provided in Appendix Table A3. We find no evidence that conditions within the treatment group triggered the invention of hybrid corn seeds.

[23] The magnitude of the treatment effect is large. However, in absolute terms the reduction in the urbanization rate is smaller. For the average county, the estimates imply the urbanization rate fell by 8.4 percentage points over the decade.

density, manufacturing productivity and other forces (Michaels et al., 2012; Gollin et al., 2016). The estimates in column 2 of Table 4 continue to show a significant negative relationship between hybrid corn seed and urbanization within the treatment group. We also find that counties with higher population density, unemployment and fertility rates in 1930 tended to have somewhat lower urbanization rates post 1935. The interaction between manufacturing productivity in 1930 and the Post dummy variable is statistically insignificant.[24]

To dig deeper into what underpins the reduction in urbanization, we inspect how the hybrid technology affected the employment share of different sectors. Column 3 of Table 4 reports estimates of equation (4) using the agricultural employment share as the dependent variable. We find that after 1935 the agricultural employment share significantly increased within the treatment group, relative to the counterfactual. Economically, the coefficient implies that the agricultural employment share increased by around 5.5%.

In column 4 of Table 4 table we find the hybrid technology had quite different effects on the manufacturing sector. Following the invention of hybrid corn seed the manufacturing share of employment fell by approximately 32% within the average treated county. While the magnitude of the coefficient estimate is large, manufacturing only accounts for 5% of employment in the average county. Hence, the manufacturing share of employment contracted by approximately 1.6 percentage points.

Columns 5 and 6 of Table 4 report estimates for the service sector. In column 5 the interaction coefficient is again negative and highly statistically significant indicating that hybrid corn seed caused a reduction in the tradable services employment share. Economically, the average treatment effect indicates that after 1935 the tradable service sector employment share fell by approximately 5.7% within the treatment group, relative to the counterfactual. The effect of the technology on non-tradable services is quite different. In column 6 the interaction coefficient is statistically insignificant. This finding is consistent with prior theoretical contributions (Hornbeck and Keskin, 2015). Specifically, agricul-

[24]The results are unchanged when we interact the county-level corn suitability index with the Post dummy.

tural productivity shocks create demand for non-tradable services which offsets the effect of rising land and labor costs.

Considering the large urbanization effects we find, an obvious question is to what extent did the hybrid technology increase agricultural output. The evidence in column 7 of Table 4 suggests that hybrid corn seeds caused approximately a 7.5% in agricultural output per capita. Moreover, the raw data show substantial increases in the quantity of corn produced within the treatment group. The average corn-growing county produced approximately 23% more bushels of corn (equivalent to 104,000 bushels) in 1940 compared to 1930. Econometric tests reported in Appendix Table A3 show that corn output increased by approximately 21% following the invention of hybrid corn seed. These large output effects are consistent with the technology greatly increasing agricultural labor demand. Hence, the large effect of the hybrid technology on urbanization and the structure of counties' economies appears plausible.

Finally, we study the long-run effects of hybrid corn seed on urbanization to examine whether the effects were long lasting. We estimate equation (4) using a sample containing urbanization rates in each county in 1930 and 1990. The results in column 7 of Table 4 show that the rate of urbanization is approximately 22% lower in the treatment group in 1990, relative to the counterfactual. Hence, the technology had a persistent impact. A potential explanation for why this is the case is that the hybrid technology generated agglomeration economies which led to a permanently more agrarian and less urbanized economy through time within treated counties.

[Insert Table 5]

The theoretical apparatus that underlies our tests suggests that the patterns in urbanization and structural change are driven by rising land and labor prices. To inspect whether this is the case we exploit the fact the Census provides data on land values and manufacturing wages.[25] Intuitively, one would expect the productivity shock to trigger an increase in both land and labor costs.

[25]Unfortunately, the Census data do not contain information on agricultural wages or land rental costs.

The results in columns 1 and 2 of Table 5 provide evidence that supports these mechanisms. Estimates in column 1 show that land prices significantly increased within the treatment group following the invention of hybrid corn seed. Economically, the effect is equivalent to a 29% increase. In addition, we find that manufacturing wages increased by approximately 38% in treated counties. The coefficient estimate is statistically significant at the 5% level.

Together these findings paint a consistent picture. Improvements in crop productivity raise agricultural output which increases agricultural labor demand to process the additional output. The increase in agricultural labor demand leads to higher wages which attracts workers from other sectors resulting in a decline in their employment share. As workers migrate from cities in search of employment in agriculture, the rate of urbanization decreases and the structure of the economy becomes more agrarian.

4.4 Robustness Checks

To ensure our findings are not driven by confounding factors, we conduct a number of sensitivity checks.

We begin by examining the robustness of our results to alternative definitions of treatment status. We first experiment by assigning a county to the treatment group based upon corn-planting patterns in 1930. We generate a dummy variable, $T1930_c$, which equals 1 if corn is planted in county c in 1930, 0 otherwise, and interact it with the Post dummy variable. The results in Panel A of Table 6 remain similar to before.

[Insert Table 6]

Given the treatment and control groups are not evenly distributed across the US (see Figure 4), we also follow the approach used by Marden (2015). This method constructs a continuous treatment variable using the normalized ratio of the suitability index for corn to the suitability index for wheat in each county. We define this suitability ratio as SCW_c and interact it with the Post dummy variable. Estimates of equation (4) using this approach are presented in Panel B of Table 6. Again, our results are comparable to the baseline findings.

[Insert Table 7]

Previous research by Hornbeck (2012) has shown that the Dust Bowl conditions between 1935 and 1938 reduced agricultural productivity and triggered a shift from agriculture to manufacturing in affected areas. If corn tends to be grown in areas that were less affected by the Dust Bowl, our results may be driven by the control group reacting to the environmental catastrophe, thereby biasing the counterfactual, and generating spuriously large average treatment effects. We control for the potential confounding effect of the Dust Bowl using county-specific erosion levels reported by Hornbeck (2012). As the dust storms began in 1935 we set the erosion variable to 0 for observations from 1930. The estimation results reported in column 1 of Table 7 show that the hybrid technology continues to cause a significant decrease in urbanization within the treatment group despite this change.

Part of the government's response to the Great Depression aimed to stimulate the economy by providing subsidies to agriculture through the Agricultural Adjustment Act of 1933 (AAA). Subsidy payments may have created incentives for individuals to remain in agriculture rather than migrate to cities. To rule out these concerns we use information on government support through the AAA in each county provided by Fishback (2005). Controlling for AAA support in column 2 of Table 7 has little impact on our main results.

Our research design assumes that improvements in crop productivity were exclusively driven by developments within the corn sector. However, increases in crop productivity among other cereal crops could potentially bias the coefficient estimates. To deal with this problem we append equation (4) with controls for barley, soybean and wheat yield per acre in each county.[26] In addition, we include an interaction between the share of acres planted with corn in 1930 and the post dummy variable to capture potential differential effects arising due to initial corn planting patterns. The results in column 3 of Table 7 remain very similar to before.

[26]Corn, wheat and barley account for the majority of cereal crop acreage.

To capture agricultural productivity change more generally we calculate the average farm size in each county. Previous evidence shows that firm size and productivity are highly correlated (Helpman et al., 2008). Including the farm size variable has little impact on the results reported in column 4 of Table 7.

The sample window spans a period of time when the western US was developing rapidly. However, most western counties are part of the control group. A danger is that the negative urbanization effect is actually driven by urbanization increasing through time in western counties. We therefore exclude counties from western states from the sample and repeat the estimations.[27] The results in column 5 of Table 7 are robust to this change.

Next, we conduct falsification tests to verify that our findings are driven by the improvement in crop productivity within the corn sector. To implement these tests we construct county-level suitability indexes for barley and wheat using information provided by the FAO-GAEZ database. We then generate a wheat (barley) dummy variable which equals 1 if the suitability index for cultivating wheat (barley) is greater than or equal to 70, 0 otherwise, and interact the wheat/barley indicator with the Post dummy variable and re-estimate equation (4). Intuitively, these tests show whether the changes in urbanization were driven by developments in other agricultural sectors. If so, it raises the question of whether our findings are spurious. However, the coefficients on the placebo interaction terms are statistically insignificant in column 6 of Table 7 whereas the T_c-Post coefficient remains negative and statistically significant. Hence, the observed decrease in urbanization was indeed driven by developments specific to the corn sector.

In an open economy setting it is conceivable that there may be spillovers to the control group if these economies lie in close proximity to counties suited to growing corn. We therefore calculate the distance between the centroid of each control group county and the nearest treatment group county's centroid and interact the distance variable with the Post indicator. Column 7 of Table 7 demonstrates that the key findings are robust to this change.

[27]The western states are Arizona, California, Colorado, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington and Wyoming.

[Insert Table 9]

Table 9 presents the corresponding tests for the agricultural, manufacturing, and service sectors' employment shares. Across all of the sensitivity checks we continue to find very similar results compared to the baseline estimates.

A concern is that our findings capture trends in urbanization that predate the sample period. If so, then our findings are spurious. We therefore investigate whether there were parallel pre-trends in urbanization within the treatment and control groups during the period 1920 to 1930. To implement this test, we estimate equation

$$urban_{ct} = \alpha_c + \beta D1930_t + \varphi T_c * D1930_t + \varepsilon_{ct}, \tag{5}$$

where $urban_{ct}$ is the urbanization rate in county c at time t; $D1930_t$ is a dummy variable equal to 1 if an observation is from 1930, 0 for 1920; T_c is a dummy variable equal to 1 if the county is in the treatment group, 0 otherwise; α_c represent county fixed effects; and ε_{ct} is the error term.

The results of this test are presented in Appendix Table A4. Reassuringly, we find the interaction coefficient to be statistically insignificant. Hence, the results are not driven by differential trends in urbanization rates between the treatment and control groups prior to the invention of hybrid corn seed.

5 Conclusions

We study how innovations trigger changes in urbanization rates by causing the growth and contraction of industries. Using a natural experiment, we obtain robust evidence that hybrid corn, a productivity-enhancing agricultural technology, provokes a significant reallocation of employment from manufacturing and services towards agriculture. Urbanization rates fall in the face of increasing crop productivity as workers migrate from urban to rural areas.

As in any quasi-experiment, a question is to what extent our findings generalize to other contexts. During our sample period US GDP per capita (in 2010 US$) was approximately $6,700 and agriculture accounted for 26% of employment. This is not dramatically different from conditions in some major developing countries such as China ($7,300, 18%), Colombia ($7,600, 16%), Peru ($6,200, 28%) and Thailand ($6,100, 33%) today. The agricultural production function is also similar between our sample period and contemporary developing countries. For example, farmers rely on rain-fed irrigation and pesticides and chemical fertilizers are rare. The similarities between our sample period and the situation in developing countries suggests that similar results may arise through the adoption of hybrid corn in developing countries today. Indeed, evidence shows that the adoption of hybrid corn and other more productive seed varieties has provoked effects similar to our findings in India, Kenya and Zambia (Foster and Rosenzweig, 2004; Smale and Mason, 2013; Smale and Olwande, 2014).

Our research has some noteworthy implications for policymakers. Firms' profit incentives often drive innovation (Acemoglu and Linn, 2004). Hybrid corn presents a case where private firms lacked incentives to conduct research because the expected returns were negative (Wallace and Bressman, 1923, 1949). Almost all the basic research into the technology was done by academic researchers (Wallace, 1955). Without public funding for universities and agricultural research stations provided by the Morrill Act of 1862 and the Harding administration during the early 1920s it seems unlikely that hybrid corn would have received significant attention despite its transformative potential. This demonstrates some of the externalities of academic research and the importance of public funding in aiding the discovery of technologies the private sector would otherwise fail to invest in, despite the high societal returns.

Innovation often depends on collaboration between a network of researchers (Liyanage, 1994). Hybrid corn was developed through the efforts of researchers at the University of Illinois, the Connecticut Experiment Station and a limited number of other universities. Knowledge spillovers between these researchers was essential in producing advances. Our work demonstrates an important role for policymakers in creating an environment for research clusters to form.

Finally, breakthrough technologies are often prized because of their transformative economic effects. For example, they create jobs in new industries. Our findings illustrate that major innovations can also produce unintended consequences such as the decline of existing industries and regions. Policymakers must develop solutions to manage these transitions, especially in adversely affected areas that often struggle to adapt as the industries they rely on contract. There is also a role for policymakers in managing housing and migratory pressures in areas where the new technology is concentrated.

References

Acemoglu, D. and Linn, J. (2004). Market size in innovation: Theory and evidence from the pharmaceutical industry. *Quarterly Journal of Economics*, 119(3):1049–1090.

Ashraf, Q. and Galor, O. (2013). Dynamics and stagnation in the malthusian epoch. *American Economic Review*, 101(5):2003–2041.

Audretsch, D. and Klepper, S., editors (2000). *Innovation, Evolution of Industry and Economic Growth*, volume Three volume set. Edward Elgar Publishing.

Aw, B., Roberts, M., and Xu, D. (2008). R&d investments, exporting, and the evolution of firm productivity. *American Economic Review*, 98(2):451–456.

Bartel, A., Ichniowski, C., and Shaw, K. (2007). How does information technology affect productivity? plant-level comparisons of product innovation, process improvement, and worker skills. *Quarterly Journal of Economics*, 122(4):1721–1758.

Bertrand, M., Duflo, E., and Mullainathan, S. (2004). How much should we trust difference-in-difference estimates? *Quarterly Journal of Economics*, 119(1):249–275.

Boothby, D., Dufour, A., and Tang, J. (2010). Technology adoption, training and productivity performance. *Research Policy*, 39(5):650–661.

Boustan, L., Fishback, P., and Kantor, S. (2010). The effect of internal migration on local labor markets: American cities during the great depression. *Journal of Labor Economics*, 28(4):719–746.

Brynjolfsson, E. and Hitt, L. (2003). Computing productivity: Firm-level evidence. *Review of Economics and Statistics*, 85(4):493–808.

Bustos, P., Caprettini, B., and Ponticelli, J. (2016). Agricultural productivity and structural transformation: Evidence from brazil. *American Economic Review*, 106(6):1320–1365.

Clark, C. (1940). *The Conditions of Economic Progress*. London: MacMillan and Co.

Coad, A., Segarra, A., and Teruel, M. (2016). Innovation and firm growth: Does firm age play a role? *Research Policy*, 45:387–400.

Conley, T. (1999). Gmm estimation with cross sectional dependence. *Journal of Econometrics*, 92(1):1–45.

Costinot, A., Donaldson, D., and Smith, C. (2015). Evolving comparative advantage and the impact of climate change in agricultural markets: Evidence from 1.7 million fields around the earth. *Journal of Political Economy*.

Criscuolo, P., Nicolaou, N., and Salter, A. (2012). The elixir (or burden) of youth? exploring differences in innovation between start-ups and established firms. *Research Policy*, 41:319–333.

Crow, J. (1998). 90 years ago: The beginning of hybrid maise. *Genetics*, 148:923–928.

Danisewicz, P., McGowan, D., Onali, E., and Schaeck, K. (2017). Debt priority structure, market discipline, and bank conduct. *Review of Financial Studies*, forthcoming.

Darwin, C. (1859). *The Origin of Species*. New York: P.F. Collier & Son.

Darwin, C. (1876). *The Effects of Cross- and Self-Fertilization in the Vegetable Kingdom.* New York: P.F. Collier & Son.

Dixon, R. (1980). Hybrid corn revisited. *Econometrica*, 48(6):1451–1461.

Dolata, U. (2009). Technological innovations and sectoral change: Transformative capacity, adaptability and patterns of change: An analytical framework. *Research Policy*, 38(6):1066–1076.

Fagerberg, J. and Verspagen, B. (2002). Technology-gaps, innovation-diffusion and transformation: An evolutionary interpretation. *Research Policy*, 31(8-9):1291–1304.

Field, A. (1978). Sectoral shifts in antebellum massachusetts: A reconsideration. *Exploration in Economic History*, 15:146–171.

Fischer, G., Nachtergaele, F., Prieler, S., Teixeira, E., Toth, G., van Velthuizen, H., Verelst, L., and Wiberg, D. (2012). Global agro-ecological zones model documentation. *IIASA Documentation Paper.*

Fishback, P. (2005). The new deal and retail sales. *Journal of Economic History*, 65(1):36–71.

Foster, A. and Rosenzweig, M. (2004). Agricultural productivity growth, rural economic diversity, and economic reforms: India 1970-2000. *Economic Development and Cultural Change*, 52(3):509–542.

Gollin, D., Jedwab, R., and Vollrath, D. (2016). Urbanization with and without industrialization. *Journal of Economic Growth*, 21(1):35–70.

Gollin, D., Parente, S., and Rogerson, R. (2002). The role of agriculture in development. *American Economic Review Papers and Proceedings*, 92(2):160–164.

Griliches, Z. (1957). Hybrid corn: An exploration in the economics of technological change. *Econometrica*, 25(4):501–522.

Hall, B. (2011). Innovation and productivity. *NBER Working Paper 17178*.

Hall, B., Mairesse, J., and Mohnen, P. (2011). *Measuring the Returns to R&D*. Amsterdam and New York.

Helpman, E., Melitz, M., and Rubenstein, Y. (2008). Estimating trade flows: Trading partners and trading volumes. *Quarterly Journal of Economics*, 123(2):441–487.

Holm, J. and Ostergaard, C. (2015). Regional employment growth, shocks and regional industrial resilience: A quantitative analysis of the danish ict sector. *Regional Studies*, 49(1):95–112.

Hornbeck, R. (2012). The enduring impact of the american dust bowl: Short- and long-run adjustments to environmental catastrophe. *American Economic Review*, 102(4):1477–1507.

Hornbeck, R. and Keskin, P. (2014). The historically evolving impact of the ogallala aquifer: Agricultural adaptation to groundwater and drought. *American Economic Journal: Applied Economics*, 6(1):190–219.

Hornbeck, R. and Keskin, P. (2015). Does agriculture generate local economic spillovers? short-run and long-run evidence from the ogallala aquifer. *American Economic Journal: Economic Policy*, 7(2):192–213.

Jedwab, R. (2013). Urbanization without structural transformation: Evidence from consumption cities in africa. *mimeo*.

Johnson, I. (1957). *The Role of the Experiment Stations in Basic Research relating to Corn Breeding*. Washington, DC: American Seed Trade Association.

Kuznets, S. (1957). Quantitative aspects of the economic growth of nations ii. industrial distribution of national product and labor force. *Economic Development and Cultural Change*, 5:1–111.

Kuznets, S. (1972). Innovation and adjustments in economic growth. *Swedish Journal of Economics*, 74:431–451.

Liyanage, S. (1994). Breeding innovation clusters through collaborative research networks. *Technovation*, 15(9):553–567.

Marden, S. (2015). The agricultural roots of industrial development: Rural savings and industrialisation in reform era china. *Mimeo*.

Matsuyama, K. (1992). Agricultural productivity, comparative advantage, and economic growth. *Journal of Economic Theory*, 58(2):317–334.

Michaels, G., Rauch, F., and Redding, S. (2012). Urbanization and structural transformation. *Quarterly Journal of Economics*, 127:535–586.

Mokyr, J. (1976). *Industrialization in the Low Counties, 1795-1850.* Yale Univ. Press, New Haven, CT.

Nunn, N. and Qian, N. (2011). The potato's contribution to population and urbanization: Evidence from a historical experiment. *Quarterly Journal of Economics*, 126(2):593–650.

Quatraro, F. (2009). Innovation, structural change and productivity growth: Evidence form italian regions 1980-2003. *Cambridge Journal of Economics*, 33(5):1001–1022.

Roback, J. (1982). Wages, rents, and the quality of life. *Journal of Political Economy*, 90(6):1257–1278.

Rosenstein-Rodan, P. (1943). Problems of industrialisation of eastern and south-eastern europe. *Economic Journal*, 53:202–211.

Russell, W. (1974). Comparative performance for maize hybrids representing different eras of maize breeding. Proceedings of the 29th Annual Corn Sorghum Research Conference.

Schumpeter, J. (1942). *Capitalism, Socialism and Democracy.* London: Unwin.

Shull, G. (1908). The composition of a field of maize. *American Breeders Association*, 4:296–301.

Shull, G. (1909). A pure line method of corn breeding. *American Breeders Association*, 5:51–59.

Shull, G. (1910). Hybridization methods in corn breeding. *American Breeders Magazine*, 1:98–107.

Shull, G. (1911). The genotypes of maize. *American Nature*, 45:243–252.

Skinner, J. and Staiger, D. (2007). *Technology Adoption from Hybrid Corn to Beta-Blockers*. NBER: University of Chicago Press.

Smale, M. and Mason, N. (2013). Hybrid seed, income, and inequality among smallholder maize farmers in zambia. *IAPRI Working Paper 72*.

Smale, M. and Olwande, K. (2014). Demand for maize hybrids and hybrid change on smallholder farms in kenya. *Argricultural Economics*, 45:1–12.

Syverson, C. (2011). What determines productivity? *Journal of Economic Literature*, 49(2):326–365.

Troyer, A. (2009). Development of hybrid corn and the seed corn industry. *Handbook of Maize*, pages 87–114.

USDA (1962). Agricultural statistics 1962. *mimeo*.

Wallace, H. (1955). *Public and Private Contributions to Hybrid Corn - Past and Future*. Washington, DC: American Seed Trade Association.

Wallace, H. and Bressman, E. (1923). *Corn and Corn Growing*. Des Moines, IA: Wallace Publishing Co.

Wallace, H. and Bressman, E. (1949). *Corn and Corn Growing 5th Edition*. New York: John Wiley & Sons.

Weber, M. and Rohracher, H. (2012). Legitimizing research, technology and innovation policies for transformative change. *Research Policy*, 61(6):1037–1047.

Wright, G. (1979). Cheap labor and southern textiles before 1880. *Journal of Economic History*, 39:655–680.

Tables

Table 1: Summary statistics

Variable	Obs	Mean	Std. Dev.	Min	Max	Level	Source
A: Agricultural Productivity Data Set							
Yield	4,070	23.44	15.60	0.10	67.90	Industry	NASS
Acres planted	4,070	10.75	1.21	4.94	12.99	Industry	NASS
Output	4,070	13.51	1.64	5.70	16.32	Industry	NASS
Hybrid share	4,070	0.08	0.11	0.00	31.00	Annual	USDA (1962)
B: Urbanization Data Set							
AG share	6,114	46.22	46.02	1	99	County	Authors' calculations
M share	6,114	4.91	12.27	0	93	County	Authors' calculations
S^T share	6,114	37.34	5.21	0	99	County	Authors' calculations
S^N share	6,114	11.53	4.76	0	45	County	Authors' calculations
Urbanization	6,114	21.31	24.25	0.01	99	County	US Census
T_c	6,114	0.09	0.29	0	1	County	Authors' calculations
Land price	6,114	20.44	486.53	0	3544	County	Authors' calculations
Wages	6,114	9.14	25.18	0	473	County	Authors' calculations
Population density$_{1930}$ * Post	6,114	31.45	125.80	0	2,963	County	US Census
Unemployment rate$_{1930}$ * Post	6,114	4.21	2.56	0	13.89	State	Authors' calculations
Manufacturing productivity$_{1930}$ * Post	6,114	3.49	5.01	0	73.34	County	US Census
Fertility$_{1930}$ * Post	6,114	21.86	117.94	0	99.50	County	US Census
Corn acreage$_{1930}$ * Post	6,114	0.03	0.17	0	1.00	County	Authors' Calculations
Manufacturing establishments	6,114	0.09	25.38	0	105	County	US Census
Dust Bowl	6,114	8.74	22.87	0	99	County	Hornbeck (2012)
AAA	6,114	32.30	59.14	0	677	County	Fishback (2005)
Average farm size	6,114	15.01	53.72	0.01	874.44	County	Authors' calculations
Barley yield	6,114	3.25	8.90	0	60.00	County	NASS
Soybean yield	6,114	1.66	4.87	0	27.70	County	NASS
Wheat yield	6,114	5.50	8.42	0	42.90	County	NASS
Manufacturing wages	6,114	22.65	35.58	0.01	473.36	County	Authors' calculations
Land price	6,114	6.65	14.52	0.52	760.39	County	Authors' calculations

Notes: Acres and output are measured in natural logarithms. AG share, M share, S share, urbanization and the unemployment rate are measured in percent. T_c is a dummy variable equal to 1 if a county is in the treatment group, 0 otherwise. Soybean, wheat and barley yield are measured in bushels per acre. Population density is population per square mile. Manufacturing productivity is the ratio of manufactured goods value to manufacturing employees. Fertility is measured as the net birth rate per 1,000 population. Manufacturing establishments is the number of manufacturing establishments per 1,000 population. Dust Bowl is the share of acres eroded by the Dust Bowl in percent. AAA is the $ value of payments per farmer through the New Deal program. Average farm size is the mean number of acres per farm.

Table 2: Corn Productivity and Hybrid Seeds

	1	2	3	4	5	6
Sample period:	1930-1940	1930-1940	1930-1940	1930-1940	1930-1940	1909-1975
Corn	0.4914***	0.2777***	0.4851***		0.6212***	0.3641***
	(18.97)	(8.03)	(18.79)		(22.73)	(10.27)
Post	0.0297		-0.0960	-0.0963		0.1637***
	(0.68)		(-0.32)	(-1.54)		(3.28)
Corn * Post	0.1721***		0.1838***	0.1832***		0.4233***
	(4.89)		(5.11)	(5.06)		(6.06)
Corn * D1931		-0.0824				
		(-1.55)				
Corn * D1932		0.0498				
		(1.41)				
Corn * D1933		-0.0565				
		(-1.00)				
Corn * D1934		-0.0007				
		(-0.01)				
Corn * D1935		0.4412***				
		(9.91)				
Corn * D1936		0.1067**				
		(2.19)				
Corn * D1937		0.1461***				
		(3.35)				
Corn * D1938		0.2608***				
		(7.19)				
Corn * D1939		0.1651***				
		(3.90)				
Corn * D1940		0.2602***				
		(6.52)				
Population density$_{1930}$ * Post			-0.0001	-0.0001	-0.0001	0.0002
			(-1.45)	(-1.44)	(-1.44)	(1.33)
Unemployment rate$_{1930}$ * Post			-0.0359	-0.0361	-0.0397	0.1755***
			(-0.43)	(-0.43)	(-0.47)	(4.80)
Manufacturing productivity$_{1930}$ * Post			-0.0151**	-0.0151**	-0.0150***	0.0122***
			(-5.52)	(-5.21)	(-5.21)	(3.96)
Fertility$_{1930}$ * Post			-0.0217	-0.0217	-0.0217	-0.0347**
			(-1.53)	(-1.52)	(-1.53)	(-2.47)
Corn * Hybrid share					0.0113***	
					(3.96)	
County FE	Yes	Yes	Yes	No	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
County * Industry FE	No	No	No	Yes	No	No
Observations	4,070	4,070	4,070	4,070	4,070	22,050
R^2	0.80	0.80	0.80	0.80	0.80	0.71

Notes: The dependent variable in all regressions is crop productivity, measured in yield per acre (ln). The standard errors are clustered at the county level and the corresponding t-statistics are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels respectively.

Table 3: Comparability of the Treatment and Control Groups

Variable	Difference	Std. Error	t-statistic
Urbanization	0.0362	0.0463	0.78
AG share	-0.1243	0.1206	-1.03
M share	0.1072	0.1116	0.96
S^{T} share	0.0279	0.0255	1.09
S^{N} share	0.0194	0.0192	1.01
Population density	0.1134	0.3356	0.35
State unemployment rate	0.1599	0.1280	1.25
Manufacturing productivity	0.2899	0.1932	1.50
Fertility	0.7858	0.6443	1.22
Manufacturing establishments	-0.2613	0.4779	-0.55
Manufacturing wages	-1.1890	1.1125	-1.07
Land price	0.1957	0.4236	0.46

Notes: This table reports the mean difference between the level of variable y in the treatment and control group during 1930 and the associated standard error and t-statistic. AG share denotes the agricultural employment share, M share is the manufacturing employment share, S^{T} share is the tradable service sector emploment share, S^{N} share is the non-tradable service sector emploment share. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels.

Table 4: Urbanization and Structural Change Results

	1	2	3	4	5	6	7	8
Dependent variable:	Urban	Urban	AG share	M share	S^T share	S^N share	Output	Urban
Post	0.8396***	0.6266**	-0.1802***	-0.1514***	-0.2348***	0.0578***	0.0012	0.5590
	(12.65)	(2.49)	(-301.28)	(-138.98)	(-96.64)	(3.51)	(0.53)	(1.33)
T_c * Post	-0.3955**	-0.5802***	0.0535**	-0.3844***	-0.0589***	-0.0281	0.0719***	-0.2449***
	(-2.51)	(-3.31)	(2.03)	(-3.19)	(-2.82)	(-1.37)	(10.25)	(-4.38)
Population density$_{1930}$ * Post		-0.0007**	-0.0002	-0.0052***	-0.0000	0.0000	-0.0000	0.0000
		(-2.21)	(-0.95)	(-7.16)	(-1.16)	(0.70)	(-0.82)	(1.21)
Unemployment rate$_{1930}$ * Post		-0.1126*	0.0431***	0.0114	-0.2053***	0.0054***	0.0010***	0.3813***
		(-1.73)	(2.75)	(0.45)	(-65.22)	(2.76)	(3.26)	(3.79)
Manufacturing productivity$_{1930}$ * Post		-0.0057	-0.0592***	-1.1014***	-0.0019	0.0002	0.0005**	-0.2240***
		(-0.49)	(-18.32)	(-265.01)	(-0.76)	(0.17)	(2.52)	(-10.89)
Fertility$_{1930}$ * Post		-0.1376***	0.0092	-0.1177***	-0.0072	-0.0055*	0.0007	0.4561***
		(-3.67)	(0.79)	(-4.37)	(-0.85)	(-1.67)	(0.90)	(6.85)
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6,114	6,114	6,114	6,114	6,114	6,114	6,114	6,114
R^2	0.05	0.06	0.99	0.99	0.91	0.01	0.14	0.06

Notes: Urban is the urbanization rate, AG share denotes the agricultural employment share, M share is the manufacturing employment share, S^T share is the tradable service sector emploment share, S^N share is the non-tradable service sector emploment share, Output is agricultural output per capita. In all columns the dependent variable is measured in natural logarithms. In columns 1 to 7 the sample period includes the years 1930 and 1940. In column 8 the sample period includes the years 1930 and 1990. We use the approach developed by Conley (1999) and used by Ashraf and Galor (2013) to correct for spatial correlation in the standard errors and report the corresponding t-statistics are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels respectively.

Table 5: Urbanization and Structural Change Results

	1	2
Dependent variable:	Land Price	Wages
Post	-0.1820***	-0.1026***
	(-10.03)	(-7.02)
T_c * Post	0.2578***	0.3228**
	(6.57)	(2.35)
Population density$_{1930}$ * Post	-0.0006***	0.0013***
	(-7.45)	(5.83)
Unemployment rate$_{1930}$ * Post	-0.0248	-0.0265
	(-0.48)	(-0.78)
Manufacturing productivity$_{1930}$ * Post	0.0196**	-0.1783***
	(2.52)	(-20.17)
Fertility$_{1930}$ * Post	0.0059	-0.0252
	(0.26)	(-0.78)
County FE	Yes	Yes
Observations	6,114	6,114
R^2	0.62	0.19

Notes: Land Price is the price per square mile of land in natural logarithms. Wages is the ratio of total manufacturing wages to the number of manufacturing employees in natural logarithms. We use the approach developed by Conley (1999) and used by Ashraf and Galor (2013) to correct for spatial correlation in the standard errors and report the corresponding t-statistics are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels respectively.

Table 6: Alternative Treatment Status Indicators

	1	2	3	4	5
Dependent variable	Urban	AG share	M share	S^T share	S^N share
A: Pre-treatment planting patterns					
Post	0.6348**	-0.1802***	-0.1512***	-0.2348***	0.0584***
	(2.52)	(-301.34)	(-139.43)	(-96.93)	(3.53)
$T1930_c$ * Post	-0.4712***	0.0610**	-0.2451**	-0.0641***	-0.0196
	(-2.62)	(2.04)	(-2.31)	(-3.17)	(-0.79)
Control variables	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes
Observations	6,114	6,114	6,114	6,114	6,114
R^2	0.06	0.99	0.99	0.91	0.01
B: Marden (2015)					
Post	0.6256**	-0.1802***	-0.1514***	-0.2348***	0.0577***
	(2.49)	(-301.29)	(-139.00)	(-96.66)	(3.50)
SCW * Post	-0.6034***	0.0592**	-0.3876***	-0.0619***	-0.0292
	(-3.22)	(2.00)	(-3.05)	(-2.83)	(-1.33)
Control variables	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes
Observations	6,114	6,114	6,114	6,114	6,114
R^2	0.06	0.99	0.99	0.91	0.01

Notes: Urban is the urbanization rate, AG share denotes the agricultural employment share, M share is the manufacturing employment share, S share is the service sector emploment share. In all columns the dependent variable is measured in natural logarithms. The control variables are Population density$_{1930}$ * Post, Unemployment rate$_{1930}$ * Post, Manufacturing productivity$_{1930}$ * Post, and Fertility$_{1930}$ * Post. We use the approach developed by Conley (1999) and used by Ashraf and Galor (2013) to correct for spatial correlation in the standard errors and report the corresponding t-statistics are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels respectively.

Table 7: Urbanization Robustness Tests

	1	2	3	4	5	6	7
Sample period: 1930 - 1940							
Post	0.7068***	0.6884***	0.6288**	1.4368*	0.6597**	0.4965*	0.5531**
	(2.80)	(2.76)	(2.49)	(1.68)	(2.40)	(1.91)	(2.06)
T_c * Post	-0.3763**	-0.3740**	-0.3504**	-0.3598**	-0.3796**	-0.3953**	-0.2814**
	(-2.36)	(-2.34)	(-2.19)	(-2.25)	(-2.35)	(-2.48)	(-2.47)
Population density$_{1930}$ * Post	-0.0007**	-0.0007**	-0.0007**	-0.0007**	-0.0006**	-0.0007**	-0.0007**
	(-2.35)	(-2.35)	(-2.24)	(-2.24)	(-1.97)	(-2.29)	(-2.21)
Unemployment rate$_{1930}$ * Post	-0.1098*	-0.1107*	-0.1181*	-0.1095*	-0.1244*	-0.1263*	-0.1230*
	(-1.69)	(-1.70)	(-1.80)	(-1.68)	(-1.77)	(-1.93)	(-1.87)
Manufacturing productivity$_{1930}$ * Post	-0.0035	-0.0037	-0.0037	-0.0027	0.0016	-0.0041	-0.0025
	(-0.31)	(-0.32)	(-0.33)	(-0.24)	(0.13)	(-0.36)	(-0.22)
Fertility$_{1930}$ * Post	-0.1407***	-0.1407***	-0.1408***	-0.1409***	-0.1742***	-0.1402***	-0.1443***
	(-3.76)	(-3.76)	(-3.76)	(-3.77)	(-4.56)	(-3.75)	(-3.90)
Dust bowl	-0.1175						
	(-0.62)						
AAA		-0.0080					
		(-0.10)					
Barley yield			0.0234				
			(1.02)				
Soybean yield			-0.0214				
			(-1.33)				
Wheat yield			0.0013				
			(0.06)				
Corn acreage$_{1930}$ * Post			0.5221*				
			(1.80)				
Average farm size				0.1662			
				(0.93)			
Wheat * Post						0.2548	
						(0.62)	
Barley * Post						0.0105	
						(0.03)	
Distance * Post							0.0360
							(1.45)
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6,114	6,114	6,114	6,114	5,300	6,114	6,114
R^2	0.06	0.06	0.06	0.06	0.06	0.06	0.06

Notes: The dependent variable is the urbanization rate, measured in natural logarithms. We use the approach developed by Conley (1999) and used by Ashraf and Galor (2013) to correct for spatial correlation in the standard errors and report the corresponding t-statistics are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels respectively.

Table 8: Structural Change Robustness Tests

	1	2	3	4	5	6	7
Panel A: AG share							
Post	-0.1796***	-0.1784***	-0.1805***	-0.1755***	-0.1821***	-0.1799***	-0.1880***
	(-301.96)	(-315.60)	(-300.71)	(-43.87)	(-320.72)	(-282.95)	(-298.83)
T_c * Post	0.0426**	0.1620***	0.1088***	0.0697**	0.1234***	0.0630**	0.1344***
	(2.06)	(4.93)	(3.52)	(2.29)	(4.43)	(2.11)	(4.61)
Population density$_{1930}$ * Post	-0.0002	-0.0004*	-0.0001	-0.0002	0.0001	-0.0002	-0.0001
	(-0.98)	(-1.79)	(-0.83)	(-0.87)	(0.45)	(-0.97)	(-0.71)
Unemployment rate$_{1930}$ * Post	0.0444***	0.0355**	0.0376**	0.0435***	0.0139	0.0452***	0.0283*
	(2.88)	(2.46)	(2.41)	(2.80)	(0.98)	(2.87)	(1.82)
Manufacturing productivity$_{1930}$ * Post	-0.0589***	-0.0577***	-0.0589***	-0.0586***	-0.0536***	-0.0592***	-0.0581***
	(-18.50)	(-18.68)	(-18.23)	(-18.03)	(-15.95)	(-18.36)	(-18.42)
Fertility$_{1930}$ * Post	0.0102	0.0119	0.0083	0.0091	-0.0082	0.0090	0.0060
	(0.89)	(1.08)	(0.72)	(0.79)	(-0.67)	(0.78)	(0.53)
Dust bowl	-0.3715***						
	(-9.26)						
AAA		-0.3301***					
		(-14.37)					
Barley yield			0.0215***				
			(5.76)				
Soybean yield			-0.0086***				
			(-3.63)				
Wheat yield			-0.0005				
			(-0.12)				
Corn acreage$_{1930}$ * Post			-0.1178***				
			(-3.02)				
Average farm size				0.1042			
				(1.18)			
Wheat * Post						-0.1200	
						(-1.46)	
Barley * Post						0.0849	
						(1.04)	
Distance * Post							0.0398***
							(6.43)
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6,114	6,114	6,114	6,114	5,300	6,114	6,114
R^2	.0.99	0.99	0.99	0.99	0.99	0.99	0.99

Table 9: Structural Change Robustness Tests

	1	2	3	4	5	6	7
Panel B: M share							
Post	-0.1520***	-0.1546***	-0.1519***	-0.1264***	-0.1526***	-0.1518***	-0.1492***
	(-138.38)	(-142.15)	(-138.21)	(-21.69)	(-126.67)	(-129.51)	(-116.78)
T_c * Post	-0.3789***	-0.5098***	-0.3191***	-0.3376***	-0.4381***	-0.3880***	-0.4635***
	(-3.16)	(-4.29)	(-2.63)	(-2.80)	(-3.65)	(-3.22)	(-3.79)
Population density$_{1930}$ * Post	-0.0052***	-0.0049***	-0.0052***	-0.0051***	-0.0049***	-0.0052***	-0.0053***
	(-7.13)	(-6.97)	(-7.15)	(-7.11)	(-6.96)	(-7.14)	(-7.11)
Unemployment rate$_{1930}$ * Post	0.0100	0.0241	0.0027	0.0144	-0.0127	0.0079	0.0312
	(0.40)	(0.99)	(0.11)	(0.57)	(-0.48)	(0.31)	(1.20)
Manufacturing productivity$_{1930}$ * Post	-1.1017***	-1.1037***	-1.1010***	-1.0982***	-1.1094***	-1.1016***	-1.1028***
	(-265.32)	(-266.99)	(-265.00)	(-264.50)	(-238.80)	(-264.49)	(-268.95)
Fertility$_{1930}$ * Post	-0.1188***	-0.1227***	-0.1192***	-0.1179***	-0.1717***	-0.1179***	-0.1134***
	(-4.44)	(-4.58)	(-4.44)	(-4.39)	(-5.15)	(-4.37)	(-4.23)
Dust bowl	0.3961***						
	(3.42)						
AAA		0.6082***					
		(12.19)					
Barley yield			0.0359***				
			(3.78)				
Soybean yield			-0.0294***				
			(-3.07)				
Wheat yield			-0.0567***				
			(-6.75)				
Corn acreage$_{1930}$ * Post			-0.3644				
			(-1.04)				
Average farm size				0.5527***			
				(4.34)			
Wheat * Post						-0.1677	
						(-0.82)	
Barley * Post						0.2375	
						(1.16)	
Distance * Post							-0.0548***
							(-3.92)
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6,114	6,114	6,114	6,114	5,300	6,114	6,114
R^2	.0.99	0.99	0.99	0.99	0.99	0.99	0.99

Table 8 Cont'd: Structural Change Robustness Tests

	1	2	3	4	5	6	7
Panel C: S^T share							
Post	-0.2354***	-0.2368***	-0.2349***	-0.2051***	-0.2331***	-0.2341***	-0.2318***
	(-92.57)	(-89.05)	(-97.22)	(-10.72)	(-104.94)	(-80.39)	(-97.63)
T_c * Post	-0.0584***	-0.0666***	-0.0575***	-0.0533**	-0.0703***	-0.0584***	-0.0697***
	(-2.80)	(-3.18)	(-2.65)	(-2.51)	(-3.42)	(-2.79)	(-3.27)
Population density$_{1930}$ * Post	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000**	-0.0000	-0.0000*
	(-1.14)	(-0.35)	(-1.12)	(-0.60)	(-2.01)	(-1.22)	(-1.67)
Unemployment rate$_{1930}$ * Post	-0.2054***	-0.2045***	-0.2055***	-0.2049***	-0.2019***	-0.2048***	-0.2026***
	(-64.81)	(-66.22)	(-66.57)	(-64.92)	(-88.90)	(-58.46)	(-74.18)
Manufacturing productivity$_{1930}$ * Post	-0.0020	-0.0021	-0.0019	-0.0016	-0.0037	-0.0019	-0.0021
	(-0.78)	(-0.82)	(-0.76)	(-0.61)	(-1.49)	(-0.75)	(-0.85)
Fertility$_{1930}$ * Post	-0.0074	-0.0076	-0.0073	-0.0073	-0.0044	-0.0072	-0.0066
	(-0.86)	(-0.89)	(-0.85)	(-0.85)	(-0.51)	(-0.85)	(-0.78)
Dust bowl	0.0383						
	(1.34)						
AAA		0.0374***					
		(3.23)					
Barley yield			0.0008				
			(0.21)				
Soybean yield			-0.0005				
			(-0.30)				
Wheat yield			-0.0064**				
			(-2.03)				
Corn acreage$_{1930}$ * Post			-0.0005				
			(-0.02)				
Average farm size				0.0657			
				(1.54)			
Wheat * Post						0.0237	
						(0.55)	
Barley * Post						-0.0337	
						(-0.79)	
Distance * Post							-0.0075*
							(-1.70)
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6,114	6,114	6,114	6,114	5,300	6,114	6,114
R^2	0.91	0.91	0.91	0.91	0.93	0.91	0.91

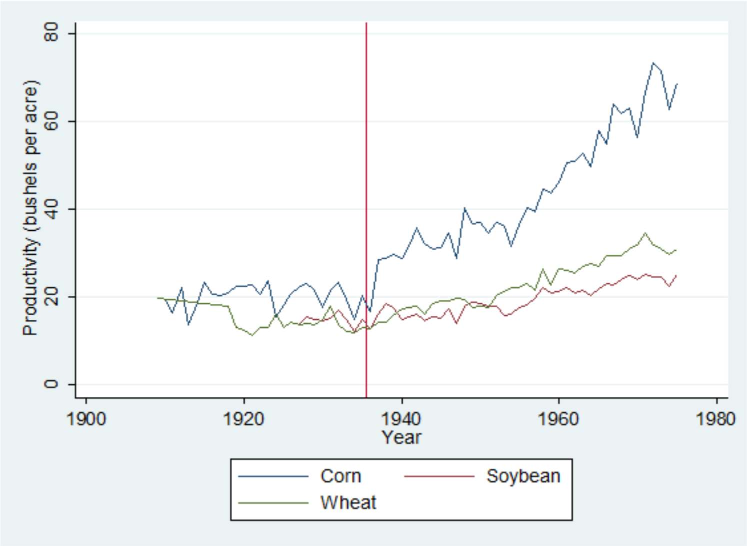
Table 8 Cont'd: Structural Change Robustness Tests

	1	2	3	4	5	6	7
Panel D: S^N share							
Post	0.0673***	0.0457***	0.0525***	0.2875	0.0656***	0.0601***	0.0584**
	(3.77)	(3.03)	(3.34)	(0.97)	(3.37)	(3.35)	(2.48)
T_c * Post	-0.0289	-0.0328	-0.0222	-0.0238	-0.0228	-0.0279	-0.0284
	(-1.40)	(-1.57)	(-1.07)	(-1.11)	(-1.11)	(-1.38)	(-1.35)
Population density$_{1930}$ * Post	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	(0.69)	(0.99)	(0.82)	(0.92)	(0.72)	(0.71)	(0.68)
Unemployment rate$_{1930}$ * Post	0.0056***	0.0059***	0.0046**	0.0057***	0.0063***	0.0056***	0.0054**
	(2.83)	(2.90)	(2.53)	(2.78)	(2.78)	(2.88)	(2.22)
Manufacturing productivity$_{1930}$ * Post	0.0003	0.0001	0.0003	0.0005	0.0005	0.0002	0.0002
	(0.21)	(0.11)	(0.21)	(0.44)	(0.36)	(0.17)	(0.17)
Fertility$_{1930}$ * Post	-0.0054	-0.0057*	-0.0057*	-0.0055*	-0.0069**	-0.0055*	-0.0055*
	(-1.62)	(-1.70)	(-1.68)	(-1.67)	(-2.05)	(-1.68)	(-1.70)
Dust bowl	-0.0560***						
	(-3.84)						
AAA		0.0227*					
		(1.83)					
Barley yield			0.0032				
			(1.11)				
Soybean yield			-0.0015				
			(-0.78)				
Wheat yield			-0.0013				
			(-0.73)				
Corn acreage$_{1930}$ * Post			0.0327				
			(0.95)				
Average farm size				0.0508			
				(0.78)			
Wheat * Post						-0.0136	
						(-1.53)	
Barley * Post						0.0104	
						(1.17)	
Distance * Post							-0.0002
							(-0.05)
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6,114	6,114	6,114	6,114	5,300	6,114	6,114
R^2	0.01	0.01	0.01	0.01	0.01	0.01	0.01

Notes: The dependent variable is measured in natural logarithms across all columns and panels of the table. We use the approach developed by Conley (1999) and used by Ashraf and Galor (2013) to correct for spatial correlation in the standard errors and report the corresponding t-statistics are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels respectively.

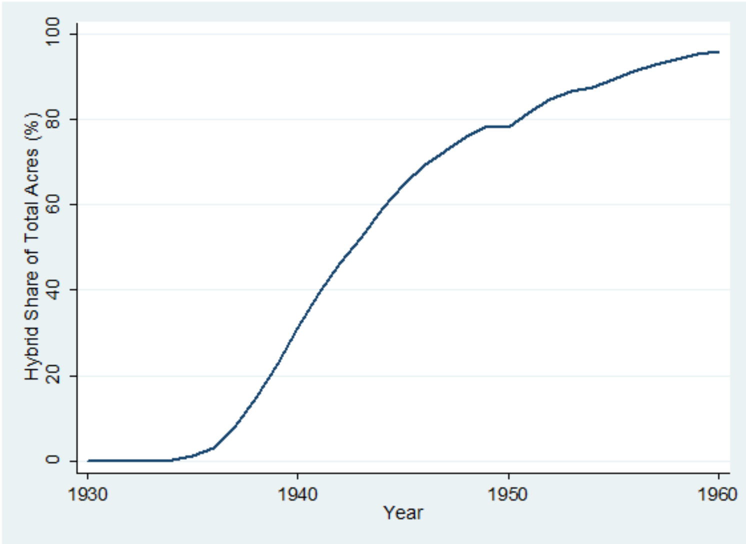
Figures

Figure 1: Historic Corn Productivity



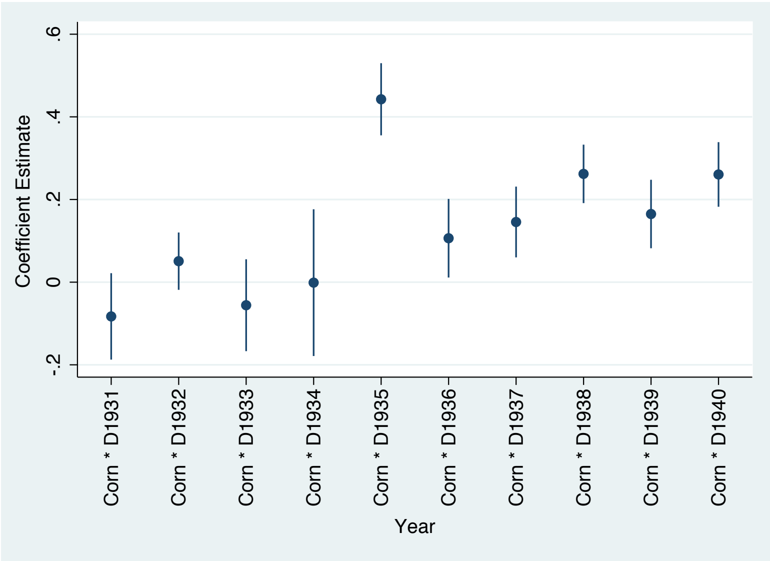
Notes: This figure plots annual productivity per acre in the US corn, soybean and wheat industries industry between 1909 and 1975. The vertical red line denotes 1935, the year when hybrid corn seed first became commercially available.

Figure 2: Percentage of Corn Acres Planted with Hybrid Seeds



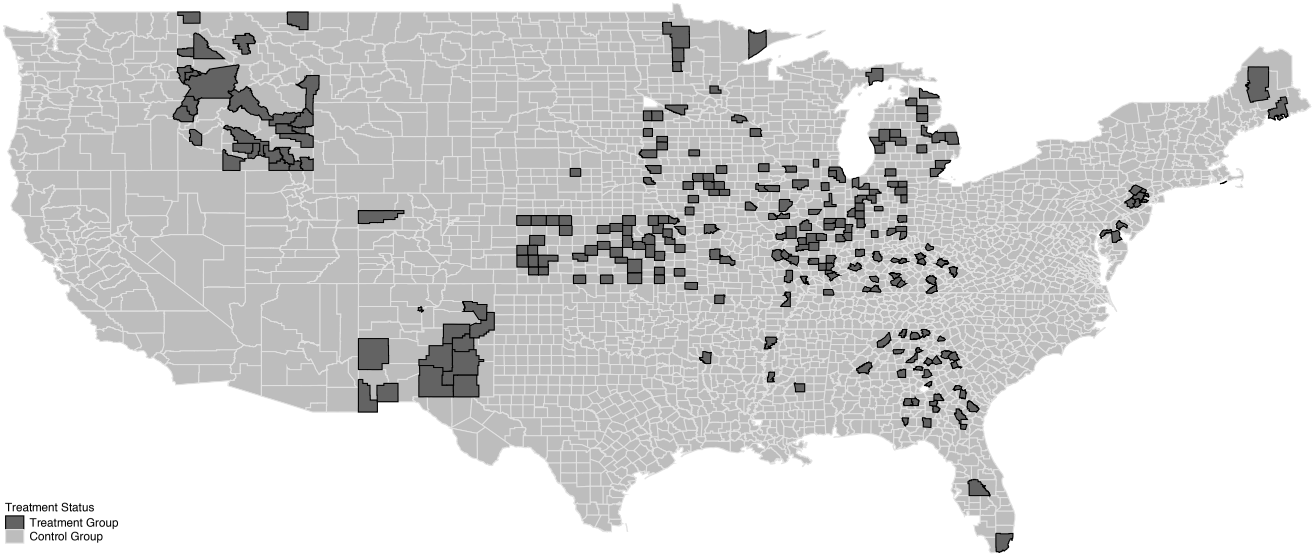
Notes: This figure plots the annual percentage of corn acres planted using hybrid corn seed between 1930 and 1960. Data is taken from the USDA (1962) *Agricultural Statistics*, Table 46, page 41.

Figure 3: Hybrid Seeds and Corn Productivity



Notes: This figure reports the interaction coefficients between the corn and year dummies from equation (2). The straight lines denote 95% confidence intervals.

Figure 4: Location of the Treatment and Control Groups



Notes: This figure shows the counties that are in the treatment (dark shading) and control (light shading) groups.

Appendix

Appendix A: Choice of the Treatment Group Threshold

Table A1: Treatment Group Threshold Choice

	1
Dependent variable: p_{it}	
T_c	0.7083***
	(13.01)
$T50_c$	0.3746***
	(19.53)
$T30_c$	0.0102
	(0.45)
Year FE	Yes
Observations	6,114

Notes: Coefficient estimates represent marginal effects evaluated at the mean. The sample includes observations from the years 1930 and 1940. We use the approach used by Conley (1999) and Ashraf and Galor (2013) to correct for spatial correlation in the standard errors and report the corresponding t-statistics are reported in parentheses. *** indicates statistical significance at the 1% level.

Appendix B: Timing of the Invention of Hybrid Corn

One may question whether the timing of the invention of hybrid corn was endogenous with respect to conditions within corn-growing counties. We investigate this hypothesis using the methodology outlined by Danisewicz et al. (2017). Specifically, if the advent of hybrid corn was endogenous to conditions within corn-producing counties, then factors such as corn productivity, planted acreage, and corn prices should be able to predict when hybrid corn was invented. We therefore estimated the equation

$$hybrid_t = \alpha + \beta_1 yield_t + \beta_2 acres_t + \beta_3 price_t + \varepsilon_t, \tag{6}$$

where $hybrid_t$ is a dummy variable equal to 1 if hybrid corn is available in year t, 0 otherwise; $yield_t$ is the mean yield per acre in corn-growing counties during year t; $acres_t$ is the mean number of acres of corn planted in corn-growing counties during year t; $price_t$ is the mean final price per bushel of corn in year t; ε_t is the error term. We use annual data for the period 1866 to 1935.

Intuitively, one would expect to find significant coefficient estimates in equation (6) if growing conditions in corn-producing counties affected the timing of the development of hybrid corn. Insignificant coefficient estimates imply that the development of hybrid corn was not endogenous to conditions within corn-growing counties. The results of this test are provided in Appendix Table A3. All coefficient estimates are statistically insignificant.

Table A2: Hybrid Corn Timing

	1
Yield	-0.0663
	(-0.91)
Acres	0.0162
	(0.94)
Price	0.0070
	(0.53)
Observations	70
R^2	0.01

Hence, econometric tests show that the timing of the development of hybrid corn was not endogenous to conditions in corn growing counties. Indeed the historical literature surrounding the invention of hybrid corn seeds is explicit that this process was driven by A) academic curiosity, and B) breakthroughs were essentially random due to researchers' lack of understanding of genetics (Crow, 1998).

Appendix C: Output Effects

Table A3 reports estimates of equation (1) using acres planted and output (total bushels produced in each county-year) as the dependent variable. In column 1 the Corn-Post interaction coefficient is statistically insignificant indicating that the number of acres planted with corn did not change substantially due to the invention of hybrid corn seed. However, when we use output as the dependent variable in column 2 the interaction coefficient is equal to 0.1916 and is statistically significant at the 5% level. This implies that the quantity of corn produced increased by approximately 21%. This is consistent with the increase in yield observed in Table 2 and the insignificant effect of the hybrid technology on acres planted.

Table A3: Output Effects

	1	2
Dependent variable:	Acres planted	Output
Corn	0.0597	0.5346***
	(0.58)	(4.82)
Post	1.1007***	0.6231
	(2.81)	(1.29)
Corn * Post	-0.0110	0.1916**
	(-0.09)	(2.41)
Corn acreage$_{1930}$ * Post	0.1591**	-0.1213
	(1.98)	(-1.34)
Population density$_{1930}$ * Post	0.0001	0.0000
	(1.01)	(0.03)
Unemployment rate$_{1930}$ * Post	0.3378***	0.1512
	(3.07)	(1.10)
Manufacturing productivity$_{1930}$ * Post	-0.0017	-0.0093
	(-0.19)	(-1.03)
Fertility$_{1930}$ * Post	-0.0052	-0.0261
	(-0.14)	(-0.66)
County FE	Yes	Yes
Year FE	Yes	Yes
Observations	4,070	4,070
R^2	0.65	0.73

Notes: In all columns the dependent variable is measured in natural logarithms. The standard errors are clustered at the county level and the corresponding t-statistics are reported in parentheses. *** and ** indicate statistical significance at the 1% and 5% levels, respectively.

Appendix D: Pre-Trends Tests

Table A4: Pre-Treatment Trends Test

	1
Dependent variable:	Urbanization
$D1930_t$	0.0503***
	(14.00)
$T_c * D1930_t$	-0.0031
	(-0.74)
County FE	Yes
N	6,054
R^2	0.18

Notes: $D1930_t$ is a dummy variable equal to 1 for the year 1930, 0 otherwise. The sample uses data for 1920 and 1930. We use the approach developed by Conley (1999) and used by Ashraf and Galor (2013) to correct for spatial correlation in the standard errors and report the corresponding t-statistics are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels respectively.

Online Appendix

For online publication

The Online Appendix provides further insights into the exogeneity of hybrid corn with respect to the dependent variables of interest. We provide a discussion of the factors that motivated the development of hybrid corn, earlier traditional (single-cross) corn seed varieties, and describe the motivations of the farmers and researchers involved. We begin by briefly outlining how corn plants are bred.

Mating Corn Plants

Corn comprises a number of different varieties. Each variety has its own genetic properties that influence the characteristics of the plant that grows following germination. For example, its root structure, disease resistance, yield and a number of other characteristics.

To produce ears containing seeds, a corn plant must first be pollinated. If a single variety is sown in a field it may be pollinated by

1. its own pollen,
2. pollen from another corn plant belonging to the same variety,
3. a different variety.

1 and 2 lead to inbreds whereas 3 leads to a 'single-cross', that is, a cross-pollinated plant. In the manuscript we refer to single-cross seeds as traditional' seeds to eliminate jargon.

Hybrid corn is produced by mating two inbred corn plants. Researchers eventually discovered in the mid-1930s that double-cross hybrids (crossing two inbred plants and crossing that hybrid with the hybrid of two other inbred plants) produced reliably high yields (Crow, 1998; Troyer, 2009). When referring to 'hybrid corn' in the manuscript, we specifically mean double-cross hybrid corn seed.

Theoretical Breakthroughs: Darwin and Mendel

Attempts to develop high-yielding corn seed began around 1846, 84 years before the start of our sample period. Especially important are findings in two books written by Charles Darwin: *The Origin of Species* (1859) and *The Effects of Cross- and Self-Fertilization in the Vegetable Kingdom* (1876). Darwin (1859) outlined a theory that inbreeding has deleterious consequences among plants and animals. Darwin (1876) provides supportive evidence by showing across 60 controlled experimental trials that inbreeding can negatively affect offspring, resulting in plants that produce lower yields than their parents. This suggested that cross-breeding varieties and selection of seeds from plants with desirable characteristics was the best way to improve yields. Subsequent research by Gregor Mendel supported Darwin's arguments (Troyer, 2009).

Importantly, hybrid corn seed requires crossing inbred varieties. It is clear from Troyer (2009) that researchers did not anticipate such a result. This is unsurprising given Darwin's insights that inbreeding generally leads to lower yields. Early efforts therefore sought to cross-pollinate corn plants to improve yield and select seeds from plants with desirable characteristics.

Traditional Single-Cross Seeds (1846 - 1935)

From 1846 thousands of single-cross corn seed varieties were developed. Some of these single-cross varieties were subsequently used to produce hybrid corn seed, such that the hybrid corn geneplasm derives from single-cross ancestors. Below we list 12 single-cross varieties that were produced between 1846 and 1935. These are the most important single-cross varieties and account for approximately 73% of the hybrid corn geneplasm.

It is clear from Troyer (2009) that in no instance was development of any single-cross variety due to economic conditions, either during our sample period or at any point in time. Contemporary, and persistent historic economic conditions, therefore do not contaminate our inferences. The single-cross seed component of the hybrid geneplasm is therefore exogenous with respect to the dependent variables we are interested in.

Instead, the motivation appears to be primarily due to private profit incentives. For example, the development of the Reid Yellow Dent and Leaming Corn varieties stemmed from farmers looking to increase output on their own farms (Curran, 1919). Their varieties subsequently gained attention and allowed them to expand sales. For example, Jacob Leaming shipped corn seeds across the US. Reid Yellow Dent gained attention after James Reid won the Chicago World Fair corn show in 1893 and subsequently received orders from neighbors, and across the United States (Troyer, 2009). The Hershey family began to sell their corn seed but only after 50 years of experimentation.

Other single-cross varieties were developed for altruistic reasons. For example, Eugene D. Funk Sr. developed the Funk Reid variety as part of his mission to use his farm estate to benefit agriculture. He established Funk Research Acres on his property following a trip to the Vilmorin Seed Company's factory in France during a trip to Europe. This inspired him to search for improvements to corn seed (Troyer, 2009). In another example, after developing more productive Richey Lancaster seeds, the Richey family gave seeds to neighbors to use on their farms. This was not driven by economic incentives but to ensure their experimental plants were not cross-pollinated by non-Richey Lancaster varieties.

Importantly, single-cross seeds proved ineffective in raising long-run corn productivity. This is because corn reproduces sexually each year. This process randomly selects half the genes from a given plant to propagate the next generation. Consequently, desirable genes (such as yield) can be lost in subsequent generations as high- and low-yielding plants are randomly mated through wind pollination. Selection of high yielding plants was therefore not a viable long-run solution (see footnote 10). While many US farmers purchased seed from seed companies, they also relied upon seed produced by their own crop during the previous harvest for seed. These seeds were likely wind pollinated by nearby corn plants such that their high-yield traits were lost. The development of single-cross varieties therefore had little effect on corn productivity. This is evident in Figure 1 where corn productivity remains low before 1935.

Below we provide details about 12 of the most important single-cross varieties that contributed to the hybrid corn geneplasm. Using information from Troyer (2009) we report the developer of each variety, the years over which they developed the variety, discuss what genetic traits they sought to encourage, the location of the research, contribution of the seeds to the hybrid geneplasm and a summary of whether economic conditions motivated their research based on evidence reported by Troyer (2009).

Reid Yellow Dent
Developer(s): Robert Reid, James L. Reid
Year: 1846 - 1910
Discussion: The Reids selected seeds from high-yielding plants in the hope that future generations of seed would continue to produce high yields. Robert Reid began developing Reid Yellow Dent corn seeds in 1846. Over a number of years he cross pollinated Gordon Hopkins and Little Yellow corn plants that produced higher-yielding offspring. James L. Reid continued these practices between 1866 and 1910 by selecting seeds from corn plants with better agronomic traits (e.g. medium size ears, bright yellow kernel color, and kernels with solid, deep, and relatively smooth grains).
Location: Peoria, Illinois
Hybrid geneplasm share: 4%
Did the economic situation motivate research: no.

Leaming Corn
Developer(s): Jacob Leaming
Year: 1855 - 1885
Discussion: Leaming selected corn seeds from heavy, earlier ripening ears after removing weak and barren plants. **Location**: Wilmington, Ohio.
Hybrid geneplasm share: 2%
Did the economic situation motivate research: no.

Lancaster Sure Crop
Developer(s): Isaac Hershey, Noah Hershey
Year: 1860 - 1920
Discussion: Isaac Hershey and Noah Hershey developed Lancaster Sure Crop by selecting seeds from plants with medium-length, well-matured, sound firm ears with clean shanks, no mould, no silk-cut kernels (i.e. not damaged or decayed), strong roots, and larger ears. They then crossed the corn seeds with Reid Yellow Dent seeds 8 to 12 times by mixing seed sown in the field. From 1910 Noah Hershey stopped crossing and blending seeds and selected based on the uniformity of corn plants.
Location: Lancaster County, Pennsylvania
Hybrid geneplasm share: 4%
Did the economic situation motivate research: no.

Chester Leaming
Developer(s): Ezra E. Chester
Year: 1885

Discussion: Chester selected Leaming Corn plants that ripened first. He then grew these in isolation. The resulting Chester Leaming seeds subsequently became the parent variety in the first inbred corn seeds (Shamel, 1901; Holden 1948; Troyer, 2004).
Location: Champaign, Illinois
Hybrid geneplasm share: 3%
Did the economic situation motivate research: no.

Richey Lancaster
Developer(s): David Richey, Frank Richey, Fredrick Richey **Year**: 1888-1920
Discussion: The Richeys first grew corn plants from Lancaster Sure Crop seeds. They then selected seeds from plants with better germination, heavier test weight and longer ears.
Location: LaSalle, Illinois
Hybrid geneplasm share: 9%
Did the economic situation motivate research: no.

Funk Reid
Developer(s): Eugene D. Funk Sr.
Year: 1888-1920
Discussion: Funk began experiments with corn in 1892. After acquiring 2,500 bushels of corn seed from a number of US corn growers, he selected the 3,000 finest ears for breeding. Through selection Funk developed Funk Reid from Reid Yellow Dent in 1900. Disease resistant plants were primarily selected.
Location: Illinois
Hybrid geneplasm share: 3%
Did the economic situation motivate research: no. Troyer (2009) reports that Funk was searching for a method to utilize Funk farms (a 25,000 acre farm) to help agriculture (Troyer, 2009). A visit to the Vilmorin Seed Company in France during a trip to Europe sparked Funks interest in producing corn seeds. He subsequently established Funk Research Acres on his farmland.

Troyer Reid
Developer(s): David Troyer
Year: 1894 - 1936
Discussion: Troyer selected seeds by removing husks, silks and trash from the ear before storing them in a dry place during the fall. The dried seeds germinated better during the subsequent planting season.
Location: Illinois
Hybrid geneplasm share: 15%
Did the economic situation motivate research: no. Troyers' neighbors asked to buy seed from him based on word of mouth and observation of his higher yields.

Minnesota 13
Developer(s): Willet Hays
Year: 1890s
Discussion: Hays grew 100 plants per generation from a single selected plants. Of the 100, he selected the best plant and repeated the process. Selection was based on heavy, mature ears with high protein content in the grain (Shoesmith, 1910; Wallace and Bressman, 1923)
Location: Minnesota
Hybrid geneplasm share: 13%
Did the economic situation motivate research: no.

Bloody Butcher
Developer(s): Oscar Will
Year: 1891 - 1894
Discussion: Will selected the 15 to 20 earliest plants with more developed ears.
Location: Bismarck, North Dakota
Hybrid geneplasm share: 5%
Did the economic situation motivate research: no.

Iodent Reid
Developer(s): Lyman Burnett
Year: 1909 - 1922
Discussion: Burnett selected Reid Yellow Dent based on earliness and yield over a 13 year period.
Location: Iowa Experiment Station
Hybrid geneplasm share: 3%
Did the economic situation motivate research: no.

Osterland Reid
Developer(s): Henry Osterland
Year: 1910 - 1920
Discussion: Osterland selected seeds from plants with earlier maturity, longer ears and looser husks.
Location: Faulkner, Iowa
Hybrid geneplasm share: 11%
Did the economic situation motivate research: no.

Strain 176A
Developer(s): Jim Holbert
Year: 1915 - 1936
Discussion: Holbert began working as a corn breeder at the Research Acres of Funk Farms in 1915, employed on a US Department of Agriculure disease project. After completing a PhD (University of Illinois) in 1936 he became Funk Research Acres' research director. He developed Strain 176A by selecting disease-free ears from disease-free plants and based on germination tests.
Location: Funk Farms, Illinois & Arlington, Virginia **Hybrid geneplasm share**: $< 1\%$
Did the economic situation motivate research: no.

Academic Research: University and Agricultural Research Stations

Johnson (1957), Crow (1998) and Troyer (2009) report that almost all research into hybrid corn seed was conducted by researchers at agricultural research stations and universities. Whereas single-cross varieties were developed by farmers, researchers and seed companies, only academics played a major role in developing hybrid corn (Wallace, 1955).

Academic research on corn was to a large extent driven by the Morrill Act of 1862 that established Land Grant colleges across the US. Notably, academic research into corn began soon after the Morrill Act whereas previously it was mainly farmers who conducted corn seed research

(Troyer, 2009). The Morrill Act was motivated by a need to provide training in agricultural practices, science and engineering to a larger section of the US population. Delivering this training was achieved by establishing new colleges across the US. Implementation of the Morrill Act was unrelated to corn productivity.

Of special importance to our paper is one Land Grant college, the University of Illinois, located in Urbana-Champaign. Being situated in the heart of the Corn Belt meant the university became the world leader in corn breeding research (Troyer, 2009). This was in part due to its close proximity to early corn seed varieties (e.g. Burr White and Chester Leamington), a number of faculty members interested in corn (Davenport, East, Holden, Hopkins, Love, Morrow described below) who shared research interests (hybrid and inbred corn plants, improving corn quality and raising soil fertility). Troyer (2004) notes that much of the research leading to hybrid corn occurred there and at the Connecticut Experiment Station.

We review several notable academics and their contributions and motivations for their research into hybrid corn. The researchers were primarily motivated by theoretical contributions. For example, Darwin (1859) and Shull (1908, 1909, 1910, 1911) are especially important. Interactions between researchers also influenced their research agenda and experiments. For example, William Beal initiated his studies following exchanges with Charles Darwin. Eugene Davenport and Perry Holden became collaborators after Holden was hired following a reorganization of the University of Illinois Agronomy Department. Edward East was initially a chemist but became interested in applying his ideas using corn following his masters dissertation at the University of Illinois. George Shull was one of the leading theorists on hybrid corn seed and was motivated to study it to confirm his theoretical models. After visiting long-term crop trials at Rothamsted, England, George Morrow abandoned his career as a lawyer and eventually became a Professor at the University of Illinois researching corn (Troyer, 2009).

The evidence reported by Troyer (2009) decisively refutes that researchers interest in hybrid corn seed was related to the economic situation both during our sample period, and in previous years. Below we provide evidence on this for each of the key researchers involved in developing hybrid corn.

William Beal
Institution(s): University of Chicago and Michigan Agricultural College
Research outputs: First to cross-fertilize corn for the purpose of increasing yields through hybrid vigor. That is, by crossing a corn plant with its parents. Beal (1876, 1881) reports that crossed seeds produced more kernels than their parents.
Motivation for research: encouraged to experiment following a helpful reply from Charles Dar-

win to a letter Beal sent in 1877.
Did the economic situation motivate research: no.

George McClure
Institution(s): College of Agriculture at the University of Illinois
Research outputs: Studied hybrid vigor among Burr White and Chester Leaming varieties and their second year progeny from 1889. 16 of 18 crosses produced higher yields than their parents. Discovered that many second-year crossed seed varieties produced smaller plants than parents whereas this was not the case for first-year crossed seeds (i.e. crossing can only raise yield in the first generation but this disappears due to the trait being lost during sexual reproduction). Experimented with inbreeding corn plants using hand pollination.
Motivation for research: became a collaborator on a research project initiated by T.F. Hunt in 1889.
Did the economic situation motivate research: no.

George Morrow
Institution(s): College of Agriculture at the University of Illinois
Research outputs: Showed that crossing open pollinated varieties on average produce higher yields than their parents. Novel idea was to plant first generation hybrid seed each year. Seed would be grown using alternate rows of the two parents and detasseling the seed parent before pollination (Morrow and Hunt, 1889; McCluer, 1892; Morrow and Gardner, 1893, 1894).
Motivation for research: on a visit to Europe he was impressed by long-term crop trials at Rothamsted, England. After returning home he established Morrow Plots and abandoned his career as a lawyer and journalist.
Did the economic situation motivate research: no.

Eugene Davenport
Institution(s): College of Agriculture at the University of Illinois
Research outputs: Together with Perry Holden developed the first inbred corn seeds in 1898. Motivation for research: After becoming Dean of the Colloge of Agriculture he hired Perry Holden to assist in a reorganization of the College in 1896. After accepting Davenports offer, Holden and Davenport spent several days planning corn improvement projects and the production of inbred seed (Crabb, 1948; Holden, 1948; Troyer 2004). This began their research into the topic of hybrid corn.
Motivation for research: collaboration with new colleague.
Did the economic situation motivate research: no.

Perry Holden
Institution(s): College of Agriculture at the University of Illinois
Research outputs: Produced early hybrids by crossing inbred plants in 1898. Found hybrid plants were larger and had more vigor than inbred seeds. Became a manager and research director at Funk Research Acres in 1902 where intensive research into corn was taking place.
Motivation for research: Trained by William Beal and inherited his research interests.
Did the economic situation motivate research: no.

Edward M. East
Institution(s): University of Illinois, Connecticut Experiment Station, Harvard University
Research outputs: Studies of genetic composition of corn plants and observing that inbreeding reduces yield. The determinants of protein content in corn. Trained 20 PhD students who went

on to work on corn breeding. East was a leader in the field of corn research and his interactions with Shull, Love and Johnson were important for developing the hybrid concept (Troyer, 2009).
Motivation for research: originally a chemist, he became interested in applying chemical and genetic insights to corn following his M.S. dissertation and subsequent PhD thesis at the University of Illinois.
Did the economic situation motivate research: no.

George Shull
Institution(s): Cold Spring Harbor and Princeton University
Research outputs: Popularized theoretical ideas of hybrid corn seed in talks at the American Breeders Association meetings. Published influential research on the topic (Shull, 1908, 1909, 1910, 1911, 1952). None of his inbred or hybrid corn seeds were used commercially.
Motivation for research: academic curiosity, interactions with Edward M. East.
Did the economic situation motivate research: no.

Donald F. Jones
Institution(s): University of Minnesota
Research outputs: Published a theory of heterosis (Jones, 1917). Made the first double-cross hybrid (the hybrid progeny from a cross between two single-cross hybrids) in 1918.
Motivation for research: academic curiosity. Development of the double-cross hybrid was due to pure luck as he mistakenly used inbred seeds in an experiment where he was testing the properties of non-inbred plants (see below).
Did the economic situation motivate research: no.

Invention and Development of Hybrid Corn Seed

Interest in hybrid corn developed in the early 1900s. George Shull, a geneticist working at Cold Spring Harbor, and subsequently Princeton University, published a series of influential papers that outlined a theory of hybrids and how they could increase yield (Shull, 1908, 1909, 1910, 1911). Speeches by Shull at the American Breeders Association also raised awareness of hybrids' potential. Shull's work was mainly theoretical and had few practical applications, a fact he recognized (Crabb, 1948). However, Shull's theoretical research provoked a great deal of academic experimentation in the search for hybrid corn.

In 1918 Donald Jones, an academic researcher working at the Connecticut Experiment Station, was conducting experiments on crossing non-inbred seed varieties. He was unaware the Chester Leaming and Burr White seeds he used were inbred varieties (Troyer, 2009). By crossing these plants, Jones inadvertently produced the first single-cross hybrid corn seed (Jones, 1927). The single-cross hybrid Jones developed was not suitable for commercial use due to poor disease resistance (Crow, 1998). However, it demonstrated that crossing inbred varieties could improve corn productivity.

The inbred seed varieties Jones used were developed by other researchers in unrelated experiments into the effects of inbreeding. The reasons he had access to these inbred varieties are as good as random.

- Inbred Chester Leaming Seeds: Researchers at the University of Illinois (Harry Love and Edward East among others) had begun an inbreeding study in 1895 using Chester Leaming seeds. Over the next 10 years they self-fertilized the original Chester Leaming plants to produce inbred plants and seeds (Jones, 1927; Troyer, 2004). Love sent a sample of the Chester Leaming inbred seeds to East in 1905 when East worked at the Connecticut Experiment Station (Crabb, 1948; Troyer, 2004). East did not use the seeds and they remained in storage until Jones' experiment in 1918.[28]

- Inbred Burr White Seeds: The inbred Burr White variety used in Jones' experiment was developed by H.K. Hayes. Hayes replaced Edward East at the Connecticut Experiment Station in 1909. While searching for a thesis topic Hayes chose to study the inheritance of protein content in corn. He studied this question using Burr White seeds. As part of his research he inbred and selected seeds from the Burr White variety between 1909 and 1913 (Hayes, 1913; Troyer, 2004). Some of the inbred Burr White seeds remained in storage at the Connecticut Experiment Station after 1913 (Troyer, 2009).

Hence, Jones had access to the inbred seeds because of unrelated research undertaken many years previously by other researchers. There are two reasons for why despite Jones' discoveries in 1918 hybrid corn did not become commercially available until the mid-1930s.

1. Jones' hybrid seeds were single-cross hybrids. While single-cross hybrids can increase crop productivity relative to traditional seeds, they lacked disease resistance traits such that few grew to maturity (Crow, 1998). Jones' results were important because they demonstrated that crossing inbred plants is a necessary condition for hybrids (Troyer, 2009).

   Obtaining seeds that can be widely used by farmers requires refinement and substantial testing through research. Wallace (1955) reports the timeline between discovery and com-

[28] Troyer (2009) reports that Harry Love and Edward E. East shared an office while working at Cornell University. Love believed that East would be interested in the inbred Chester Leaming variety because East had been experimenting with growing inbred plants while at Cornell. However, East failed to conduct any experiments and the seeds remained in storage. This appears to be because his supervisor, Dr Hopkins (Head of the Agronomy Department, University of Illinois), decided that having observed a field of inbred corn plants grown by East in 1905 they would change research direction and focus on bulking ears of inbred plants (Troyer, 2009).

mercial availability of corn seeds was 15 years (irrespective of whether they were hybrids or single-cross varieties). Our review of varieties reported by Troyer (2009) shows that the lag between invention and commercial sale was at least 15 years for most varieties developed between 1846 and 1935.

Following Jones' breakthrough in 1918 considerably more research therefore had to be conducted before commercially viable hybrids were invented. During the early 1930s academic researchers discovered double-cross hybrid corn seeds that were both more productive and disease resistant (Crow, 1998). Double-cross hybrids became commercially available in the mid-1930s. There is no evidence that economic conditions influenced research from 1918 onwards (Troyer, 2009). Rather, numerous trial-and-error experiments were conducted before double-cross seeds were discovered (Crow, 1998).

2. Mass producing hybrid corn seed was not commercially viable before the 1930s. Henry A. Wallace, the owner of a corn seed business in Des Moines, Iowa, was aware of the hybrid theory outlined by Shull (Troyer, 2009). But he reported that the production technique was too complex and expensive to be feasible (Wallace and Bressman, 1923, 1949). Indeed, Shull recognized that the breeding method he proposed was, "to be of theoretical rather than of practical interest" (Crabb, 1948).

Hence, it was only after researchers had found new ways to produce disease-resistant hybrid corn seeds that they could be mass produced. Refinements were also necessary to ensure that hybrids could grow in diverse areas (Crow, 1998).

Exogeneity of Hybrid Corn

Throughout our review of the literature we can find no evidence that economic conditions during our sample period motivated the development of hybrid corn seed. This is the case for both the researchers who developed hybrid corn, and the farmers, researchers and seed companies that produced earlier single-cross varieties.

Rather, it is clear from the discussion above that researchers were interested in hybrid corn due to their academic interests. Edward East, Donald Jones and George Shull exemplify this academic curiosity. The breakthrough in understanding how to produce hybrids was almost entirely dependent on a fortunate set of circumstances and was unrelated to economic conditions, or any of our dependent variables. Specifically, a researcher was conducting experiments with seeds he did not realize were inbred and fulfilled the necessary requirements for producing hybrid corn. The subsequent testing of hybrid corn seed by researchers was similarly unrelated to economic conditions.

Based on all of these insights, we conclude that the introduction of hybrid corn seed is exogenous with respect to our variables of interest. Specifically, there is no evidence that shows urbanization or the share of employment in the agricultural, manufacturing, or service sectors motivated the development of hybrid corn, or affected the timing of its invention. Furthermore, the invention of hybrid corn is unrelated to economic conditions and other factors contained in the error term of our estimating equations.

Online Appendix References

Crabb, A.R. (1948). *The hybrid-corn makers-prophets of plenty.* New Brunswick, NJ: Rutgers University Press.

Crow, J. (1998). 90 years ago: The beginning of hybrid maize. *Genetics*, 148: 923-928.

Curran, W.R. (1919). Indian Corn: Genesis of Reid's Yellow Dent, *Journal of the Illinois State Historical Society (1908-1984)*, 11(4): 576-585.

Darwin, C. (1859). *The Origin of Species*. New York: P.F. Collier & Son.

Darwin, C. (1876). *The Effects of Cross- and Self-Fertilization in the Vegetable Kingdom*. New York: P.F. Collier & Son.

Hayes, H.K. (1913). Report of the plant breeder. Connecticut AES Report No. 37.

Holden, P.G. (1948). Corn breeding at the University of Illinois 1895-1900. *Archives Michigan State University*, mimeo.

Johnson, I.J. (1957). *The role of the experiment stations in basic research relating to corn breeding.* Washington, DC: American Seed Trade Association.

Jones, D.F. (1917). Dominance of linked factors as a means of accounting for heterosis. *Genetics*, 2: 466-497.

Jones, D.F. (1927). Double crossed Burr-Leaming seed corn. *Connecticut Ext. Bull*, 108.

McCluer, G.W. (1892). Corn crossing. *Illinois AES Bullentin*, 21: 83-101.

Morrow, G.E. and Gardener, F.D. (1893). Field experiments with corn, 1882. *Illinois AES Bullentin*, 25: 179-180.

Morrow, G.E. and Gardener, F.D. (1894). Field experiments with corn, 1885. *Illinois AES Bullentin*, 31: 359-360.

Morrow, G.E. and Hunt, T.F. (1889). Field experiments with corn, 1888. *Illinois AES Bullentin*, 4: 48-67.

Shamel, A.D. (1901). Seed corn and some standard varieties for Illinois. *Illinois AES. Bulletin*, 63: 29-56.

Shoesmith, V.M. (1910). *The study of corn.* New York: Orange Judd.

Shull, G.H. (1908). The composition of a field of maize. *American Breeders Association Report*, 4: 296-301.

Shull, G.H. (1909). A pure line method of corn breeding. *American Breeders Association Report*, 5: 51-59.

Shull, G.H. (1910). Hybridization methods in corn breeding. *American Breeders Magazine*, 1: 98-107.

Shull, G.H. (1911). The genotypes of maize. *American Naturalist*, 45: 234-252.

Shull, G.H. (1952). Beginnings of the heterosis concept. Pages 14-48 in J.W. Gowen (ed.) Heterosis. Ames, IA: Iowa State College Press.

Troyer, A. (2004). Champaign County, Illinois, and the origin of hybrid corn. *Plant Breeding Reviews*, 24: 42-59.

Troyer, A. (2009). Development of hybrid corn and the seed corn industry. *Handbook of Maize*: 87-114.

Wallace, H. and Bressman, E. (1923). *Corn and corn growing.* Des Moines, IA: Wallace Publishing Co.

Wallace, H. and Bressman, E. (1949). *Corn and corn growing.* 5$^{\text{th}}$ Edition revised by J.J. Newlin, E. Anderson, and E.N. Bressman. New York: John Wiley & Sons. Des Moines, IA: Wallace Publishing Co.